UCL SCHOOL OF PHARMACY
BRUNSWICK SQUARE

Molecular insights into pathogenic mechanisms in startle disease

Victoria James

Thesis submitted for the degree of Doctor of Philosophy

UCL SCHOOL OF PHARMACY
Department of Pharmacology
29-39 Brunswick Square
London WC1N 1AX
United Kingdom

This thesis describes research conducted in the Department of Pharmacology, UCL School of Pharmacy between October 2009 and September 2012 under the supervision of Professor Robert J. Harvey. I certify that the research described is original and that any parts of the work that have been conducted by collaboration are clearly indicated. I also certify that I have written all the text herein and have clearly indicated by suitable citation any part of this dissertation that has already appeared in publication.

Signature Date: 19th November 2012

Abstract

This study describes an in-depth investigation into the pathogenic mechanisms of inherited mutations that lead to disorders of inhibitory glycinergic transmission, primarily the rare human disorder known as startle disease/hyperekplexia. I also investigated mutations causing a similar startle phenotype in cows, mice and zebrafish. Using molecular genetics techniques, I identified pathogenic mutations in the genes that encode for proteins involved in glycinergic neurotransmission, specifically the postsynaptic glycine receptor (GlyR) subunits and the presynaptic glycine transporter GlyT2. Using homology modelling and other computational biology methods, I examined the structural and functional impacts of mutations on protein function, revealing key motifs and amino acids crucial for receptor and transporter activity. Using cDNA cloning and site-directed mutagenesis, I also generated expression constructs for wild-type and mutant proteins that were used in functional tests to measure the impact of pathogenic mutations on glycine receptor and transporter function. For certain animal models of startle disease, I was also able to develop diagnostic PCR tests for pathogenic mutations, which can be used to alleviate further animal suffering by preventing 'at risk matings' of carrier animals. Taken together, my findings reveal several new pathogenic mechanisms of GlyR and GlyT2 mutations in startle disease in humans and animals, revealing insights into receptor and transporter function that may be applicable to other neurological disorders.

Acknowledgments

I would like to start by thanking my supervisors, Robert Harvey and Maya Topf, not only for their mentoring, but also for their tremendous support, encouragement, ideas and advice throughout my Ph.D. They have both been fantastic to work with, and I will always be so grateful that they gave me this opportunity.

My Ph.D. and the research would not have been possible without funding from the Bloomsbury Colleges Consortium: I am also enormously thankful for this support.

The work would not have been complete without collaboration with various groups and individuals, particularly Prof. Mark Rees, Rhys Thomas and Seo-Kyung Chung at the Institute of Life Science, Swansea University; Stéphane Supplisson at Institut de Biologie de l'ENS, Paris, Max Suster at the University of Bergen, Norway; Joe Lynch and Anna Bode at the Queensland Brain Institute, University of Queensland; and Brian Pearce at UCL School of Pharmacy. I must also thank Eloisa Carta whose work preceded mine, and who also helped me find my way around the lab when I first came to the School of Pharmacy.

I would also like to thank my friends and colleagues in the Harvey and Topf labs, in particular Jennifer Gill for her endless help and advice, both professionally and personally, and Irene Farabella for her moral support and amusement.

Finally, I would like to thank my wonderful family for their consistent love, support and encouragement, and my boyfriend Matt for taking care of me through the ups and downs and being there for me. I could not have done it without them.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. The glycinergic inhibitory synapse

1.1.1. Biological roles of glycine receptor subtypes

Glycine receptors (GlyRs) are pentameric ligand-gated chloride channels that mediate inhibitory synaptic transmission in the spinal cord, brainstem, cerebellum and retina. They are members of the Cys-loop ligand-gated ion channel superfamily, of which all share a common structure and topology. Each receptor is made up of five subunits that create a membrane spanning ion-channel pore. Each subunit consists of four transmembrane (TM) domains connected by inner- and outer-membrane loops, and a large extracellular domain (ECD) that is made up of β-sheets and loops. The major adult GlyR isoform, consists of α1 and β GlyR subunits, and has a major role in the control of spinal motor reflex circuits (Fig. 1.1). Mutations in the genes encoding this GlyR subtype (*GLRA1* and *GLRB*) cause excessive startle responses, characterized by noise- or touch-induced non-epileptic seizures, excessive muscle stiffness and neonatal apnoea episodes in cattle, mice and humans (Harvey *et al* 2006, 2008). However, the biological roles of other GlyR subtypes, containing the α2, α3 and α4 subunits, are still under investigation. The embryonic/neonatal GlyR α2 subtype has previously been linked to roles in synaptogenesis (Kirsch & Betz 1998; Lévi *et al* 1998), cell fate and paracrine transmitter release (Mangin *et al* 2003) in the developing cortex (Flint *et al* 1998) and spinal cord (Scain *et al* 2010) as well as retinal photoreceptor development (Young & Cepko 2010). It was therefore surprising that *Glra2* knockout mice did not show a clear behavioural phenotype (Young-Pearse *et al* 2006). However, this was also true for initial studies on GlyR α3 subunit knockout mice, where detailed phenotypic characterisation was required to reveal subtle defects in central inflammatory pain sensitisation (Harvey *et al* 2004; Hösl *et al* 2006; Harvey *et al* 2009) and rhythmic breathing (Manzke *et al* 2010). The GlyR α4 subunit has been linked to neurotransmitter release in sympathetic neurons (Boehm *et al* 1997) and is found in dorsal root ganglion neurons and the male genital ridge in chick embryos (Harvey *et al* 2000) and cholinergic amacrine cells in mouse retina (Heinze *et al* 2007). Curiously, the gene encoding GlyR α4 is thought to be a pseudogene in humans (Simon *et al* 2004)

due to a stop codon in *GLRA4* exon 9, causing a protein truncation between membrane-spanning domains TM3 and TM4. However, this finding may need revisiting in the light of recent re-sequencing studies showing that certain genes on the X chromosome are apparently functional in some individuals but contain nonsense mutations or frameshifts in other apparently normal control subjects (Tarpey *et al* 2009).

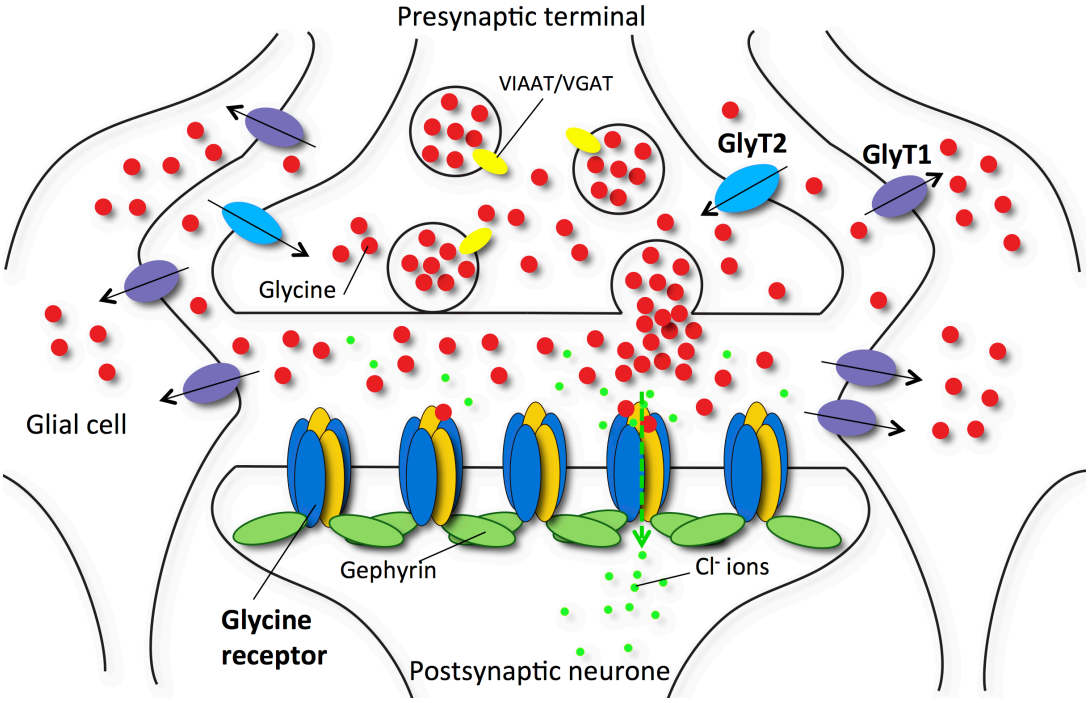

Figure 1.1. The glycinergic inhibitory synapse. Human GlyRs are each composed of two α1 subunits and three β subunits and are clustered on the membrane of the postsynaptic neurone opposite the presynaptic terminals that release glycine. GlyRs are clustered at the membrane by gephyrin. GlyT1 is located on neighbouring glial cell membranes, and is responsible for transporting glycine away from the synapse. GlyT2 is located on the presynaptic terminal membrane and is responsible for recycling glycine back into the terminal to be repackaged for release.

1.1.2. Biological roles of glycine transporter subtypes

Glycine transporters 1 and 2 (GlyT1 and GlyT2) belong to the Na^+/Cl^--dependent neurotransmitter transporter family that includes transporters for dopamine,

norepinephrine, epinephrine, serotonin and γ-aminobutyric acid (GABA) (Nelson 1998). These membrane proteins share a similar topology, having 12 transmembrane (TM) domains connected by a series of extracellular and intracellular loops, and intracellular N- and C-termini. GlyT1 and GlyT2 are encoded by two different genes, *SLC6A9* and *SLC6A5*, and although the proteins share around 50% amino acid sequence identity, they differ in function, localisation and uptake capacity. As illustrated in figure 1.1, GlyT1 terminates glycinergic neurotransmission via uptake of glycine from the synapse into surrounding glial cells, whilst GlyT2 mediates the uptake of glycine into the presynaptic terminal, providing glycine for vesicular inhibitory amino acid transporter (VIAAT)-mediated refilling of synaptic vesicles (Eulenburg *et al* 2005). These roles were revealed by the generation of GlyT1 and GlyT2 knockout mice (Gomeza *et al* 2003a,b), which show normal histology and synaptic protein expression, but show differing neuromotor defects that ultimately lead to their premature death. GlyT1 knockout mice display severe motor and respiration deficits at birth with lethargy, hypotonia and hyporesponsivity resulting from overinhibition. The pups die on the first postnatal day, indicating that GlyT1 is not essential for embryonic development but is vital for postnatal CNS functions. The reduction of overall motosensory functions in the GlyT1-deficient mice is similar to some symptoms characteristic of a group of human inherited disorders known as glycine encephalopathy or nonketotic hyperglycinemia (NKH) (Gomeza *et al* 2003a). By contrast, GlyT2 knockout mice are normal at birth, but in the second postnatal week develop a severe neuromotor phenotype characterised by spasticity, a rigid muscle tone, strong tremor and a severely impaired righting response. They die prematurely by the end of the second postnatal week as a result of being unable to feed and continued convulsions. The symptoms observed in these mice are similar to those associated with the hereditary human motor disease hyperekplexia. Cultured neurons from GlyT2 knockout mice have a significant reduction in the amplitude of miniature inhibitory currents, reflecting the reduced presynaptic glycine release in the mutant animals (Gomeza *et al* 2003b).

1.2. Defects at the glycinergic inhibitory synapse and neurological disease

1.2.1. Startle disease/Hyperekplexia: a rare and treatable neurological disorder

The startle reflex is a normal reflex elicited to a minor degree in normal newborns and infants, and consists of eye blinking, facial grimacing, flexion of the head, hunching of the shoulders and flexion of the elbows, trunk and knees causing falling without a protective reaction (Praveen *et al* 2001; Sharma *et al* 2006). Startle disease, also known as hyperekplexia, is a rare neurological disorder that causes an exaggerated startle reflex, characterised by noise- or touch-induced non-epileptic seizures, neonatal hypertonia and apnoea (suspension of breathing). Initial studies (Shiang *et al* 1993; 1995) suggested that hyperekplexia was predominantly an autosomal dominant disease, although more recent studies suggest that recessive cases are more common (Rees *et al* 2006; Chung *et al* 2010; Carta *et al* 2012). Although rare, hyperekplexia can have serious consequences such as brain damage or sudden infant death (Harvey *et al* 2008). Clonazepam, a benzodiazepine that potentiates the action of GABA at the $GABA_A$ receptor can be used to treat the hypertonia and apnoea associated with hyperekplexia. In addition, the Vigevano manoeuvre, which entails the flexing of the head and limbs towards the trunk, has also proved to be effective particularly when prolonged stiffness affects respiration (Vigevano *et al* 1989; Praveen *et al* 2001). Hyperekplexia is now recognized and detected within specialist neurology and paediatric centres worldwide (Davies *et al* 2010).

1.2.2 Mutations in the adult GlyR α1β isoform cause startle disease/hyperekplexia

The primary cause of hyperekplexia in humans is missense, nonsense, deletion and frameshift mutations in the GlyR α1 subunit gene (*GLRA1*) on chromosome 5q33.1 (Shiang *et al* 1993; 1995). The most common mutations in *GLRA1* to date are autosomal dominant missense mutations (depicted as red circles in Fig. 1.2), and typically affect residues in the second transmembrane domain, which lines the Cl^- ion channel that is essential for the GlyR function, or the extracellular TM2-TM3 loop nearby. These mutations typically lead to a decrease in agonist affinity by causing disruptions in the signal transduction pathway that links glycine binding to the gating

of the integral chloride channel. A number of recessive *GLRA1* mutations have also been identified, which are thought to either be a result of the inheritance of single alleles in consanguineous families or compound heterozygosity of different mutations (blue circles in Fig. 1.2; Harvey *et al* 2008a). In fact, it has recently been shown that recessive hyperekplexia caused by *GLRA1* mutations is more common than dominant hyperekplexia on a population basis (Chung *et al* 2010). Mutations in the GlyR α1 subunit gene also cause similar disorders in mice (Buckwalter *et al* 1994; Ryan *et al* 1994; Holland *et al* 2006) and Poll Hereford cattle (Pierce *et al* 2001) (purple circles in Fig. 1.2). Several mouse models of startle disease due to mutations in the mouse *Glra1* have been discovered. *spasmodic*, *oscillator* and *cincinnati* are spontaneous, recessive mutations in mouse *Glra1*. They are characterised by an exaggerated startle reflex in response to sudden stimuli giving rise to a state of rigidity with tremor, muscle spasms, falling and an impaired righting reflex. *spasmodic* is the result of a missense mutation leading to an A52S substitution in the mouse GlyR α1 subunit, which does not result in fatality, due to only slightly reduced agonist sensitivity (Ryan *et al* 1994). By contrast, the mutant mouse *oscillator* results in a lethal phenotype whereby the mutant mice will die by three weeks of age. In this case, a microdeletion in exon 8 results in a complete loss of GlyR function (Buckwalter *et al* 1994). The *cincinnati* mutant phenotype closely resembles that of *oscillator*, and is the result of a duplication of exon 5 in *Glra1,* ultimately leading to a frameshift mutation and loss of function (Holland *et al* 2006). *Nmf11* is the result of an N-ethyl-N-nitrosourea-induced missense mutation (N46K; Traka *et al* 2006). The substitution results in a significant reduction in agonist sensitivity (Harvey and Smart, personal communication), which results in a lethal phenotype resembling that of *oscillator* and *cincinnati*. In Poll Hereford cattle, a hyperekplexia-like disorder known as inherited congenital myoclonus (ICM) has also been characterised (Pierce *et al* 2001). ICM is an autosomal recessive disease and is reported to be characterized by hyperesthesia and myoclonic jerks of the skeletal musculature that occur in response to sensory stimuli and also spontaneously. A nonsense mutation in exon 2 of the bovine *GLRA1* was found to be responsible for this disorder, leading to a defective assembly of the expressed protein (Pierce *et al* 2001).

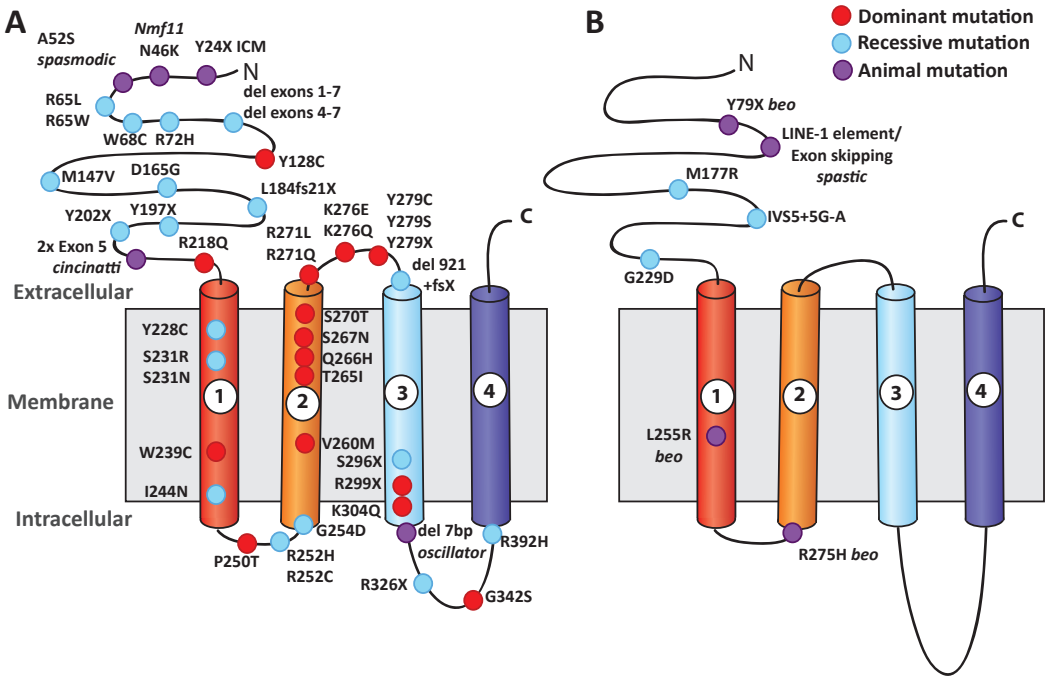

Figure 1.2. Glycine receptor (GlyR α1 and GlyR β) topology and locations of known startle disease mutations. A representation of GlyR α1 (**A**) and GlyR β (**B**) subunit topology annotated with the positions of dominant (red circles) and recessive (blue circles) mutations identified in humans and animals (purple circles) causing hyperekplexia.

To date, no mutations in any other GlyR α subunit genes (*GLRA2*, *GLRA3* and *GLRA4*) have been reported in hyperekplexia. However, mutations in other postsynaptic proteins involved in glycinergic transmission have been identified. Mutations in *GLRB* that encodes the GlyR β subunit have been reported in two families (Rees *et al* 2002; Al-Owain *et al* 2011) (blue circles in Fig. 1.2). A missense (G229D) and a splice site (IVS5+5→GA) mutation occurred together in a compound heterozygote leading to a hyperekplexia phenotype. The substitution G229D reduces agonist affinity of heteromeric α1β subunit GlyRs (Rees *et al* 2002), whereas the IVS5+5→GA splice site mutation leads to a complete loss of exon 5 in *GLRB*. Another mutation was reported very recently in nine related individuals with a classical hyperekplexia phenotype (Al-Owain *et al* 2012). This recessive missense mutation leads to the substitution of a highly conserved methionine to an arginine at residue 177 (M177R) in the extracellular domain of the GlyR β subunit (Al-Owain *et al* 2012).

Mutations in mouse *Glrb* (Kingsmore *et al* 1994) and zebrafish *glrbb* (Granato *et al* 1996; Hirata *et al* 2005; Masino & Fetcho 2005) (purple circles in Fig. 1.2) have also been reported. The spontaneous mouse mutant *spastic* is characterised by an exaggerated startle reflex in response to auditory and tactile stimuli, with severe rigidity and tremor, and impaired righting reflex (Mulhardt *et al* 1994) - a similar phenotype to the *Glra1* mouse mutants *spasmodic*, *oscillator*, *cincinnati* and *Nmf11*. The mutation responsible for the phenotype is an intronic long interspersed element 1 (LINE-1) retrotransposon that lies between exons 6 and 7. This insertion causes mis-splicing of GlyR β subunit RNAs and impairs expression of normal GlyR complexes (Kingsmore *et al* 1994; Mulhardt *et al* 1994). Zebrafish *bandoneon* (*beo*) mutants are defective in glycinergic neurotransmission due to mutations in *glrbb* encoding the zebrafish GlyR βb subunit. 24-28 hour wild-type embryos exhibit characteristic coiling movements in response to tactile stimuli, whereas *beo* mutants instead demonstrate apparent shortening of the body axis followed by an extension back to the normal body size in an 'accordion'-like fashion. This action is lethal for the mutant fish due to mechanical stress, leading to lesions within the notochord (Granato *et al* 1996). A lack of synaptic aggregation of GlyRs in *beo* mutants as a result of defects in *glrbb*, leads to a loss of reciprocal inhibition of the motor circuits between the two sides of the spinal cord. Consequently, motor neurons on both sides are activated simultaneously leading to the bilateral contraction of the axial muscles and the so-called 'accordion' phenotype (Hirata *et al* 2010).

1.2.3. Mutations in the glycine transporter GlyT2 gene are a second major cause of startle disease/hyperekplexia

Initial genetics studies established that many individuals with startle disease do not carry mutations in *GLRA1 or GLRB*. For this reason, several additional candidate genes were examined based on biological plausibility, including the clustering proteins gephyrin (Rees *et al* 2003) and collybistin (Harvey *et al* 2004). However, neither of these genes are now thought to play a major role in hyperekplexia (Harvey *et al* 2006), having been implicated in molybdenum co-factor (MOCO) deficiency (Reiss &

Hahnewald 2011) and X-linked intellectual disability (Harvey *et al* 2004; Kalscheuer *et al* 2009), respectively. However, *SLC6A5*, encoding the glycine transporter GlyT2 was considered to be an excellent candidate gene based on biological function. In 2006, Rees and co-workers identified missense, nonsense and frameshift mutations in the GlyT2 gene in individuals presenting with typical hyperekplexia symptoms: exaggerated startle response to tactile or acoustic stimuli, hypertonia and neonatal apnoea episodes. The mutations resulted in defective localisation of GlyT2, abolished [^{3}H]glycine uptake, or both. Based on the structural similarity of human GlyT2 to a bacterial leucine transporter (LeuT; Yamashita *et al* 2005), certain mutations were predicted to affect Na^+ or glycine binding residues. This was the first example of a human disorder directly related to genetic defects in a gene encoding a Na^+/Cl^--dependent neurotransmitter transporter, and suggested that neurological disorders with a postsynaptic genetic component could also have a corresponding presynaptic defect (Rees *et al* 2006).

Two important studies by Gomeza *et al* in 2003 reported the characterisation of GlyT1 and GlyT2 knockout (KO) mice. GlyT1 deficient mice ($GlyT1^{-/-}$) appear externally normal at birth but die within the first postnatal day, which is most likely due to their inability to suckle. They have a severe phenotype whereby the pups show very weak spontaneous motor activity. For example, when touched gently, wild-type or heterozygote littermates flail their limbs whereas, $GlyT1^{-/-}$ KO mice did not move (Gomeza *et al* 2003a). GlyT1 is responsible for terminating glycine neurotransmission by removing glycine from the synapse, therefore explaining their constant muscle relaxation and lack of response to touching. The phenotype of GlyT2 KO mice differs significantly from that of GlyT1. It is also severe - they gain weight slowly and die prematurely by the end of the second postnatal week. These mice appear normal at birth, but also display a complex phenotype involving rigid muscle tone, strong spontaneous tremor, spasticity, hind feet clasping and an impaired righting reflex (Gomeza *et al* 2003b). This phenotype closely resembles that of mutant mice *spasmodic*, *oscillator*, *Nmf11* and *spastic* that harbour defects in *Glra1* or *Glrb*,

suggesting that the *Slc6a5* knockout mouse is also a valuable animal model for startle disease.

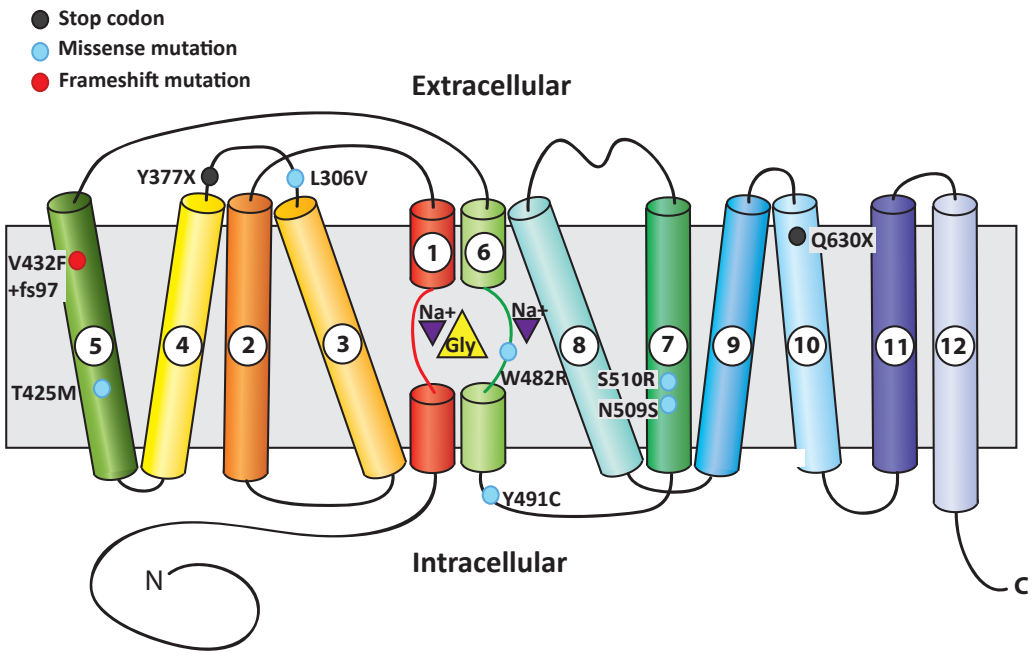

Figure 1.3. Glycine transporter GlyT2 topology and locations of known startle disease mutations. A representation of predicted GlyT2 topology within a membrane and positions of known mutations in hyperekplexia patients**.** Transmembrane domains are numbered 1-12, glycine is represented by a yellow triangle and the sodium ions by blue triangles.

1.2.4. Startle disease in cattle and dogs

GlyT2 defects also cause disorders similar to startle disease in Belgian Blue cattle (Charlier *et al* 2008) and Irish Wolfhounds (Gill *et al* 2011). Charlier *et al* (2008) identified a missense mutation in *SLC6A5* in Belgian Blue cattle with congenital muscular dystonia 2 (CMD2), where the calves exhibit symptoms highly reminiscent of hyperekplexia in humans, but typically die within a few hours of birth. CMD2 was discovered using innovative genome-wide SNP scans that localised the defective gene to a 3.61 Mb segment of bovine chromosome 29, harbouring the GlyT2 gene. The mutation T809C in exon 4 of *SLC6A5* results in a L270P substitution in the third membrane-spanning domain (TM3) of GlyT2. Further genotyping of cattle in Belgium

revealed that 2.3% of Belgian Blue cattle and 4.6% of artificial insemination sires were clinically unaffected carriers of the defective c.T809C allele. However, all 21 confirmed CMD2 cases were found to be homozygous for this change, confirming that the disorder showed recessive inheritance. The affected residue was not predicted to be involved in glycine or Na^+ binding, but the leucine was highly conserved between all vertebrates suggesting structural or functional significance. Mutation of the corresponding residue (L269P) in human GlyT2 abolished [^{3}H]glycine uptake in recombinant assays (Charlier *et al* 2008).

In 2011, Gill *et al* reported the first genetically confirmed cases of canine startle disease in Irish Wolfhounds. In a litter of seven puppies, two showed an exaggerated startle response with severe muscle stiffness and tremor in response to handling. Sequencing of *SLC6A5* revealed a 4.2 kb deletion containing exons 2 and 3 resulting in the loss of a large part of the end terminus as well as all the subsequent transmembrane domains of the protein due to a frameshift (Gill *et al* 2011). Both pups were found to be homozygous for this change, while both parents and three other pups in the litter were found to be heterozygous, confirming recessive inheritance. These findings in cattle and companion animals are not only significant for the understanding of human startle disease, but they are also key in informing future breeding strategies in order to avoid future cases and ultimately eliminate the inherited mutation from the population of animals.

1.3 Zebrafish glycinergic transmission

1.3.1 Zebrafish as a model for the study of glycinergic transmission

Significant progress has been made in the molecular genetics of zebrafish contributing to a greater understanding of the systems and mechanisms involved in locomotor network development and function (Drapeau *et al* 2002). Zebrafish have become an attractive model for the study of motor development for a number of reasons. The embryos develop very rapidly and are transparent allowing the visualization of dynamic events in live fish, such as Ca^{2+} transients and axonal outgrowth. The size of

the brain or eye, for example or length of the body of the fish can be observed at different time points and are generally very uniform between wild-type fish. Also, importantly, the behaviours of zebrafish embryos are easily assayed and well characterised. Embryos classically exhibit three defined behaviours by 36 hours post fertilization (hpf). At 17 hpf, embryos exhibit a simple slow coiling behaviour independent of stimulation, which is seen as repetitive alternating coiling of the tail. At 21 hpf the coiling becomes stronger and occurs in response to mechanosensory stimulation. Finally at 26 hpf, this coiling becomes swimming, which is initiated by mechanosensory stimulation, increasing from 7 Hz to 30 Hz at 36 hpf (Hirata *et al* 2004). In order to understand the embryonic behaviours of the zebrafish, the physiology of the embryonic neurons and muscles can be analysed using electrophysiological techniques.

Importantly, forward genetics can be applied to zebrafish, identifying genes responsible for correct motor behaviours. Large-scale mutagenesis using N-ethyl-N-nitrosourea (ENU) has generated hundreds of zebrafish mutants that are useful for dissecting synaptic transmission. Granato and co-workers (1996) performed a genetic screen detecting behavioural phenotypes for motor functions. 166 mutants were identified from this screen exhibiting embryonic motility defects, which were then categorised into 14 phenotypes, estimating to comprise defects in at least 48 genes (Granato *et al* 1996). 103 of these mutants displayed normal birefringence of muscular fibres suggesting impairments in the nervous system, neuromuscular junction or functional components of muscle such as EC coupling. These mutants were categorised according to their responses to touch: no response, normal but reduced response, vigorous but abnormal response, or simultaneous, bilateral contractions.

The zebrafish *shocked* (*sho*) mutant has a deficient escape response showing reduced spontaneous coiling at 17 hpf and failing to initiate swimming after touch at 48 hpf. They respond to tactile stimulation with a single large contraction of the trunk: the so-called 'twitch once' phenotype (Granato *et al* 1996; Cui *et al* 2005). This phenotype is

caused by missense mutations in *slc6a9*, encoding GlyT1 (Table 1.1). The ta229g allele displays the strongest phenotype and harbours a G81D missense mutation in GlyT1 (Cui *et al* 2005). The te301 allele displays a milder phenotype and harbours a C305Y missense mutation (Mongeon *et al* 2008). The mutation in the ta51e line is currently unknown. Curiously, *sho* ta229g mutants normally die within two weeks of development, but can be maintained to adulthood by careful feeding. By contrast, te301 mutants recover by around 4-5 dpf. This suggests that the G81D and C305Y substitutions show important differences in their effect on GlyT1 function, and that zebrafish GlyT1 mutants are ideal for studying both embryonic and adult synaptic transmission and behaviours. Despite these advances provided by the study of *sho* mutants, a similar ENU mutant in GlyT2 has remained elusive. However, a recent re-examination of zebrafish motility mutants that had been mapped near *slc6a5* on zebrafish chromosome 7 by Geisler and co-workers (2007) uncovered a promising lead. The macho class mutant *schlaffi* (*sla*) maps between markers z1206 and z4999, a region that harbours *slc6a5*.

Mutant	Alleles	Mutation	Phenotype and gene defect	References
shocked sho	ta51e **ta229g§** **te301†**	Unknown **G81D** **C305Y**	Embryonic lethal. d2, twitch only once, head not straight; d5, head straight, in response to touch just jumps up and falls down, then vibrates with tip of tail. Resting position sideways or upside down. **Mutations G81D and C305Y in the glycine transporter 1 gene (*slc6a9*) on chromosome 2.**	Granato *et al* 1996; Odenthal *et al* 1996; Luna *et al* 2004; Cui *et al* 2004; Cui *et al* 2005; Mongeon *et al* 2008
schlaffi sla	th239, ty112, tg230 (lost)	Unknown Unknown Unknown	Embryonic lethal. d2, reduced motility, do not rotate in chorion, normal muscle striation, fast buzzing movements. d3, jumping movements. d5, somites contracted and notochord degenerating in 50% of the embryos.	Granato *et al* 1996; Odenthal *et al* 1996

Table 1.1. Gene locations, nomenclature and mutations/phenotypes for zebrafish GlyT1 *shocked* mutant and a potential GlyT2 mutant - *schlaffi*. ‡Mutant lost, †Viable allele, §Strongest allele.

Another interesting class of zebrafish mutants is the 'accordion' mutants - *accordion* (*acc*), *zieharmonika* (*zim*), *diwanka* (*diw*), *bandoneon* (*beo*), *quetschkommode* (*que*), *bajan* (*baj*) and *expander* (*exp*), which are so-called due to the simultaneous, bilateral contractions of the axial muscles in response to tactile stimuli (Hirata *et al* 2010). Wild-type embryos at 24-28 hpf show characteristic coiling movements in response to tactile stimuli. Applying the same stimuli to the accordion mutant embryos induces the simultaneous contractions of the trunk muscles on both sides of the embryo resulting in a shortening of the body axis by 5-10%, followed by an extension back to the original size (Hirata *et al* 2004). This phenotype occurs as a result of the loss on an inhibitory mechanism that, in wild-type embryos, prevents motor neurons on one side from firing when the muscles on the other side are contracting. Additionally, it was demonstrated that this inhibitor mechanism is glycine-mediated since application of strychnine, a glycine receptor blocker, to wild-type embryos phenocopied the abnormal 'accordion' behaviour (Granato *et al* 1996; Hirata *et al* 2004).

1.3.2. The extended GlyR subunit gene families in zebrafish

Inhibitory GlyRs belong to a superfamily of ligand-gated ion channels that includes nicotinic acetylcholine receptors (nAChRs), serotonin ($5HT_3$) receptors and $GABA_A$ receptors, and are heteromultimers made up of ligand-binding α subunits and structural β subunits. The β subunits also bind gephyrin, which is a protein essential for clustering $\alpha\beta$ GlyRs at synapses. In mammals, there are four α subunit genes (*GLRA1*, *GLRA2*, *GLRA3* and *GLRA4*) and one β subunit gene (*GLRB*). Zebrafish have five α subunit genes (*glra1*, *glra2*, *glra3, glra4a* and *glra4b*) and two β subunit genes (*glrba* and *glrbb*) (Hirata *et al* 2010). Phylogenetic analysis has shown high sequence similarities of mammalian GlyR α1, GlyR α2, GlyR α3, and GlyR α4 subunits to zebrafish GlyR α1, GlyR α2, GlyR α3, and GlyR α4a/GlyR α4b subunits respectively (Imboden *et al* 2001; Hirata *et al* 2010). Zebrafish can therefore provide a very useful model for the study of the roles of the glycine receptor subunits and ultimately aid understanding of disease mechanisms of mutations that affect these genes. Table 1.2 brings together the current findings as to which glycine receptor subunit or glycine

transporter genes are responsible for mutants in the ENU screen or the effects of knockdown. For many genes the effect of knockdown is currently unknown. Similarly, there are *accordion* mutants with a currently unknown causative gene. *Expander (exp)* for example, is an interesting mutant for which the causative mutation remains unknown.

Gene	Location	Protein	Mutant/Knockdown	Key references
Glycine receptors				
glra1	Chr 14	GlyR α1	Knockdown: no phenotype reported	David-Watine *et al* 1999; Hirata *et al* 2005; McDearmid *et al* 2006
glra2	Chr 9	GlyR α2	Unknown	Hirata *et al* 2010
glra3	Chr 1	GlyR α3	Unknown	Imboden *et al* 2001b
glra4a	Chr 14	GlyR α4a	Knockdown: Disrupted rhythm-generating networks and reduced number of spinal interneurons	Imboden *et al* 2001a,b; McDearmid *et al* 2006
glra4b	Chr 5	GlyR α4b	Unknown	Imboden *et al* 2001b
glrba	Chr 1	GlyR βa	Unknown	Hirata *et al* 2005
glrbb	Chr 14	GlyR βb	*bandoneon*: touch-induced bilateral muscle contraction	Hirata *et al* 2005
Glycine transporters				
slc6a9	Chr 2	GlyT1	*shocked*: touch-induced single twitch	Higashijima *et al* 2004 Cui *et al* 2005 Mongeon *et al* 2008
slc6a5	Chr 7	GlyT2	Unknown	Higashijima *et al* 2004

Table 1.2: Known zebrafish GlyR and GlyT mutants/knockdowns and their phenotypes.

1.4. Protein structure prediction by homology modelling.

Large-scale genome sequencing projects have resulted in millions of known protein sequences from various organisms. However, protein structure determination is limited by cost, time and experimental challenges that, unfortunately, are primarily inherent to X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy. Therefore, the rate of experimental structure determination for these proteins is much slower, with structure known for only around one hundredth of those that have been characterised (Eswar 2003). However, with the help of computational methods it is possible to generate reliable structural models based on existing structures of other proteins to examine a given hypothesis. Homology modelling has been a popular tool for the last decade and is increasingly relied upon as the number

of known sequences continues to rise at a faster rate than the number of known protein structures.

Homology modelling is completed in four stages (Fig 1.4). The first stage is to identify sequence homologues with known structures that may be used as a template structure. This is done using fold recognition software and web servers such as HHPred (Söding *et al* 2005) and GenThreader (Lobley *et al* 2009), which detect the closest related structures that should be used as a template structure on which to base the homology model. An approximate sequence alignment between the target sequence and the structure is created so that a percentage sequence identity may be calculated. The second stage of homology modelling is the alignment stage. The target sequence is carefully and more accurately aligned with the selected template sequence (or sequences), taking into consideration regions of high sequence similarity, conservation within the family and secondary structure. Sequence alignment methods have come a long way since the method of visual comparison. Now many web servers are available to create these alignments, having different algorithms and different levels of accuracy. In order to perform a pairwise alignment where just two sequences are aligned, dynamic programming is used which is based on an algorithm devised by Needleman and Wunsch in 1970. A table of scores for matches and mismatches between amino acids (e.g. the BLOSUM62 matrix (Eddy 2004)) is given, along with defined penalties for insertions and deletions. The Needleman-Wunsch algorithm (Needleman & Wunsch 1970) is generally used for global alignments where the program attempts to align all of the residues/nucleotides in the entire length of the sequence, and the Smith-Waterman algorithm (Smith & Waterman 1981) is used for local alignments, which is more useful for sequences that lack overall similarity but are thought to have areas of local similarity or shared motifs. Heuristic methods which are less accurate but more efficient are employed for database searches such as BLAST (Basic Local Alignment Search Tool) (Altschul *et al* 1990), where short subsequences are taken from the query sequence and matched to a database of candidate sequences. If multiple subsequences are matched between two sequences in the same

locality then the method will go on to apply a more sensitive alignment, while the sequences with no similarity are discarded. Multiple sequence alignments are very useful for identifying regions of conservation within a group of related sequences, which is of particular importance in homology modelling when identifying important structural motifs. Multiple sequence alignments are more computationally difficult to calculate in comparison to pairwise alignments, but are becoming faster due to increasing computational power and more efficient algorithms. MUSCLE (Multiple Sequence Comparison by Log-Expectation), for example, can align, with accuracy, up to 5000 sequences of an average length of 350 amino acids in less than 10 minutes on a desktop computer (Edgar 2004). MUSCLE uses a progressive, or tree method (Feng & Doolittle 1987) where the first of the two most similar sequences are aligned by a pairwise alignment and then further sequences are added successively in the order of most to least related until all sequences in the query set have been included (Feng & Doolittle 1987).

The third stage of homology modelling is the model building phase, where models are calculated and built computationally using the alignment and the coordinates of the template structure or structures. MODELLER (Sali & Blundell 1993) is a popular computer program designed for this task in particular, although is also capable of carrying out all four steps in the process. This software builds comparative models using an alignment of the target sequence to the template sequence, the atomic coordinates of the template structure and a simple script to instruct the program (Sali & Blundell 1993; Eswar *et al* 2006). The program executes the model building by the satisfaction of spatial restraints, whereby a set of geometrical criteria help create a probability density function for each atom in the structure. The following spatial restraints are examined:

1. Homology-derived restraints on atomic distances and dihedral angles taken from the alignment with the target structure.
2. Stereochemical restraints: bond length and bond angle obtained from CHARMM-22 force field (Brooks *et al* 2009).

3. Statistical preferences for dihedral angles and non-bonded inter-atomic distances, using information from a representative set of known structures.

4. Optional manually designed restraints, such as those from NMR spectroscopy, rules of secondary structure packing, cross-linking experiments, fluorescence spectroscopy, image reconstruction from electron microscopy, site-directed mutagenesis and intuition (Eswar 2003).

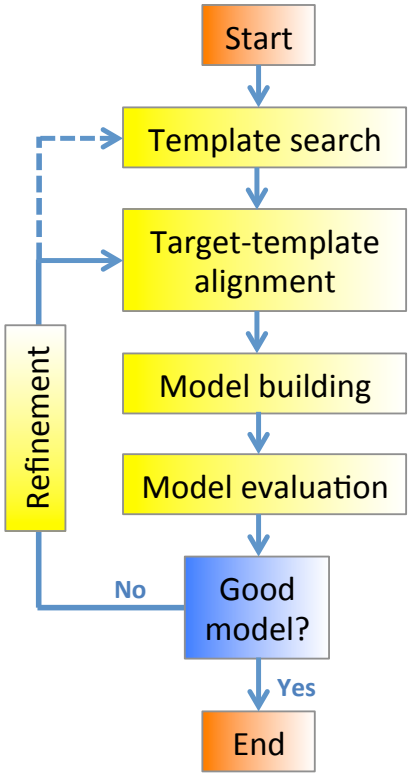

Figure 1.4. The stages of homology modelling.

The fourth and final stage of homology modelling is the model assessment stage, where the accuracy of the models created is assessed. Identifying the most accurate model from an ensemble of models is a critical step in protein structure prediction. The accuracy of a model is usually directly correlated with the sequence identity of the target protein to the template protein thus the sequence identity itself gives us the most easily quantifiable measure of the accuracy. However, this is not adequate, particularly for models built on templates with a relatively low sequence identity. Relative energy of the model can be calculated and compared to the statistical potential of the mean forces of proteins that have solved structures. These native

protein structures are found to have the lowest energies, and can be used as a consistent benchmark of how "native-like" a homology model is. These structures provide a great deal of information about the atomic interactions that dictate what makes a protein native-like. This information has been utilized to generate many and varied statistical potentials in order to assess the accuracy of homology models. The DOPE (Discrete Optimised Potential Energy) statistical potential (Shen & Sali 2006) is an atomic distance-dependent statistical potential that is derived from a sample of native structures in the protein data bank (PDB; www.pdb.org; Berman *et al* 2000). It is widely used and is incorporated within MODELLER so that a DOPE score may be calculated following model building using the same script each time. These scores allow us to predict the most accurate model or models out of many that have been built. It also scores each residue individually, which allows the determination of problematic areas such as loop regions that may require further refinement. Additional assessments can be made on a final model using web servers such as the QMEAN (Qualitative Model Energy Analysis) server (Benkert *et al* 2009), which analyses the quality of a model using evolutionary and physiochemical criteria, giving both global and individual residue quality estimates. Simple servers such as this one are extremely useful to allow experts and relative novices to realize the overall quality of a model and/or areas of the model that are unreliable and may require further refinement.

1.5. The crystal structure of LeuTAA: a bacterial homologue of Na^+/Cl^--dependent neurotransmitter transporters

Currently, there is no existing crystal structure for the vertebrate glycine transporters. Membrane proteins are notoriously difficult to purify due to their high levels of hydrophobicity, and eukaryotic proteins can also be very difficult to express. However, in 2005, Yamashita and colleagues exploited the fact that prokaryotic organisms also harbour amino acid transporters and crystallised the leucine transporter (LeuT) from *Aquifex aeolicus*. Although the overall sequence identity between the prokaryotic LeuT and eukaryotic counterparts is only 20-25%, clusters of high sequence conservation, including functionally important residues, are distributed throughout the sequences.

This makes the LeuT structure a good template for homology modelling of eukaryotic homologues, such as glycine transporters. Molecular cloning of a large family of neurotransmitter transporters, including those for GABA, glycine, serotonin and dopamine (Bröer 2006) has already provided some initial structural insights. Analysis of these sequences suggested that these transporters have 12 transmembrane regions (TMs) connected by a series of intracellular and extracellular loops, with the N- and C-termini residing in the cytoplasm. Mutagenesis and structure/function studies have identified residues involved in antagonist binding, substrate interactions, gating and ion co-ordination (Yamashita *et al* 2005; Singh *et al* 2008). In the leucine transporter, Yamashita *et al* (2005) pinpointed substrate and sodium ion binding sites (non-protein electron densities) in TM1, TM3, TM6 and TM8. In particular, two sodium ion-binding sites were identified within 6Å of the α-carbon of the bound leucine, utilizing residues in the unwound regions of TM1 and TM6 (Figure 1.5).

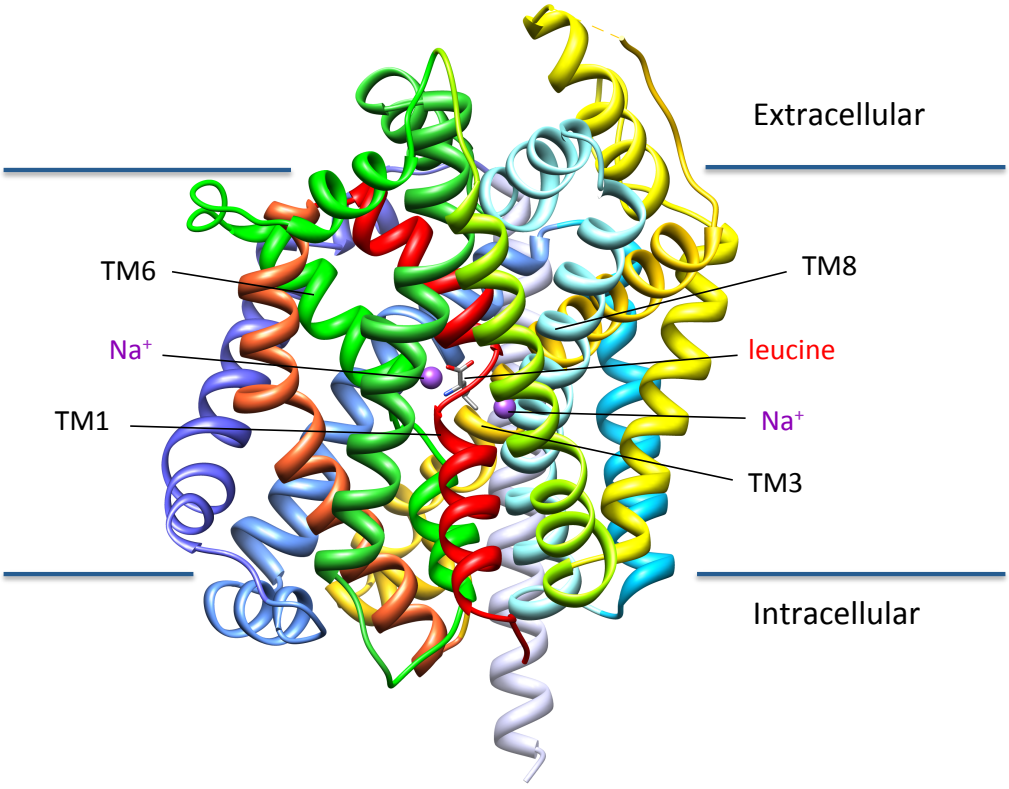

Figure 1.5. The structure of the *Aquifex aeolicus* leucine transporter. The structure was visualised using Chimera (Pettersen *et al* 2004) depicting the leucine substrate, Na^+ ions and transmembrane regions important for binding.

1.6. Crystal structure of the *C. elegans* glutamate-gated chloride channel GluCl: A helpful insight into the structures of the Cys-loop ligand-gated ion channel superfamily

Cys-loop ligand-gated ion channels are membrane-spanning, pentameric receptor proteins activated by neurotransmitters. They are thus called due to a similar motif within the extracellular domain of each subunit, where the two sides of the loop are brought together by a pair of disulphide-bridging cysteine residues. Cys-loop receptors are responsible for excitatory and inhibitory neurotransmission in the central nervous system (CNS). The Cys-loop receptor family includes the vertebrate receptors for serotonin (5-HT_3), acetylcholine (nACh), glycine (Gly) and γ-aminobutyric acid ($GABA_A$, $GABA_c$), along with a range of invertebrate (GluCl) and prokaryotic (ELIC and GLIC) Cys-loop receptors (Thompson *et al* 2010). All Cys-loop receptors share a common secondary, tertiary and quaternary structure. Each of the five subunits that make up one receptor has a large extracellular domain (ECD) at the N-terminus that is made up of β sheets and loops, and is anchored into the cell membrane by the transmembrane domain at the C-terminus that is made up of four α-helices (fig 1.6). The subunits come together to form symmetrical pentameric complexes where the second transmembrane (TM2) helix of each subunit faces the centre of the complex. Together, all five TM2 helices form a selective ion channel or pore, which passes through the centre of the complex (Fig. 1.7).

We encounter a similar problem when understanding the molecular structure and dynamics of the GlyR as we do for GlyT2, since there is currently no solved crystal structure for this receptor. The structure for the nicotinic acetylcholine receptor (nAChR) was, until recently, the closest related structure available. The initial three-dimensional electron microscopy (EM) image at a 9Å resolution was first published (Unwin 1993), and has been refined at 4Å resolution (Unwin 2005). Many predictions about the structure of the GlyR were made using this structural information, despite the relatively low sequence identity of the different nAChR subunits (GlyR α1 subunit: 19.7 to 20% and β subunit: 23.9 to 24.6%). Crystal structures of prokaryotic ligand-

gated ion channels such as *Gloeobacter violaceus* GLIC, *Erwinia chrysanthemi* ELIC and functional homologues of the amino-terminal ligand-binding domain of the nAChR (Hilf & Dutzler 2008; Bocquet *et al* 2009) have also proved to be very useful in providing structural information about the nAChR and other Cys-loop receptors, in particular, regarding the agonist binding site (Brejc *et al* 2001), but again these show limited sequence identify to GlyRs (e.g. GLIC versus GlyR α1 subunit: 21.8% and β subunit: 24.6%).

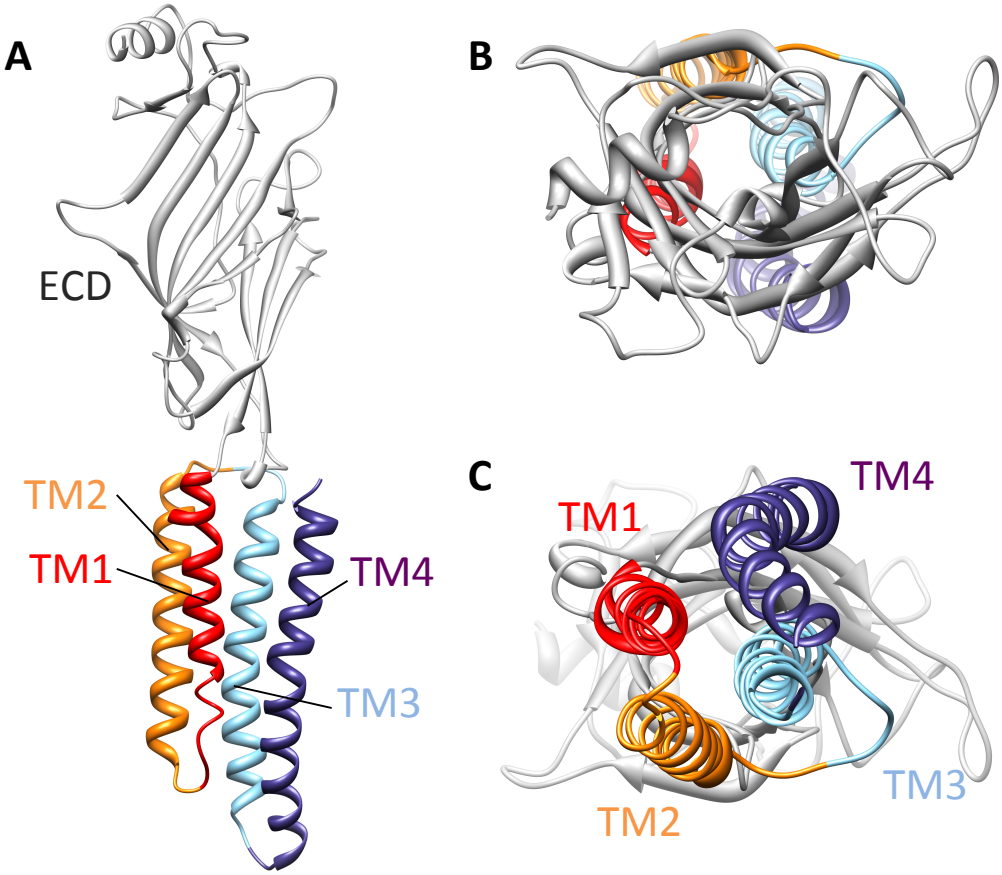

Figure 1.6. The structure of a single GluCl receptor subunit. A single subunit of GluCl represented as a ribbon diagram, showing the subunit from the side (**A**), top (**B**) and bottom (**C**). Each TM2 helix is coloured differently to make them distinguishable from each other. The ECD is coloured in grey.

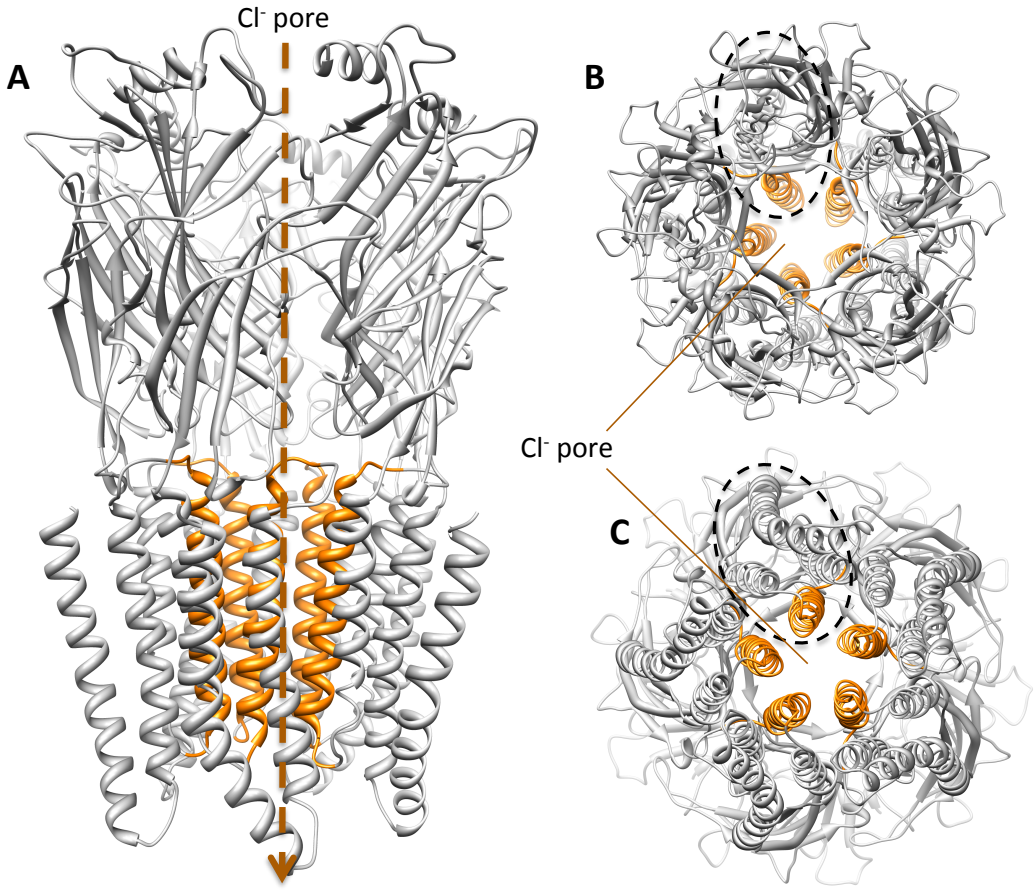

Figure 1.7. The structure of the GluCl receptor and the integral Cl^- channel. A single GluCl receptor represented as a ribbon structure, showing the pentameric quaternary structure from the side (**A**), top (**B**) and bottom (**C**). The TM2 helix of each subunit is coloured in orange to show the ion channel pore, while the other TMs and the ECD is coloured in grey.

In mid 2011, the crystal structure of a *C. elegans* glutamate-gated chloride selective ion-channel GluCl resolution was determined at 3.3Å resolution (Hibbs & Gouaux 2011). The GluCl structure is an invaluable tool for predicting the structures of various related receptors, shedding some more light on the ligand-binding site (Hibbs & Gouaux 2011) and binding of ivermectin (Duterte *et al* 2012; Lynagh & Lynch 2012), since the crystallised structure includes these molecules bound within the structure. The GluCl structure makes a significantly better template for homology modelling of the human GlyR compared to GLIC, ELIC or the *T. momorata* nAChR due to functional similarity (i.e. GluCl is an anion channel) and higher overall sequence identity - 44.4% (GlyR α1 subunit) and 38.9% (GlyR β subunit), respectively.

1.7. Aims of thesis

Until recently, dominant mutations in the gene encoding the postsynaptic GlyR α1 subunit (*GLRA1*) were considered to be the single major cause of startle disease. However, a number of novel mutations have recently been discovered in the presynaptic GlyT2 gene (*SLC6A5*) in individuals with startle disease, suggesting that GlyT2 dysfunction may represent a second major cause of the disease. In chapter 3, I aimed to understand the pathogenic mechanisms underlying GlyT2 dysfunction. Firstly, I aimed to use site-directed mutagenesis and [^{3}H]glycine uptake assays to measure glycine uptake for GlyT2 mutants. Secondly, I aimed to develop new homology models of human GlyT2 based on the structure of the bacterial leucine transporter, in order to understand how GlyT2 mutations impair transporter function.

Congenital muscular dystonia (CMD2) is an inherited disease found in cattle with a startle-disease-related phenotype. This recessive disorder can have serious economical consequences for the cattle farming industry, as well as animal welfare issues, if it remains undetected within a herd. In chapter 4, I investigate two new cases of CMD2 reported in Belgian Blue calves from the United Kingdom. I aimed to develop rapid genotyping tests for the specific genetic mutation responsible for CMD2, to be used in a preliminary population analysis study for CMD2 in cattle from the United Kingdom. An additional aim was to build a homology model of bovine GlyT2, to offer further insight into the structural and functional effects of the CMD2 mutation on GlyT2 function.

Whilst mutations in the GlyR α1 subunit and GlyT2 genes are established causes of startle disease, considerably fewer mutations have been identified the GlyR β subunit gene (*GLRB*). This is unusual, considering that these GlyR subunits are expressed together to form heteromeric receptors. In chapter 5, I aimed to determine the pathogenic mechanisms of three novel mutations in *GLRB* that were recently identified in startle disease patients, using bioinformatic analysis, homology modelling and functional assays.

In chapter 6, I aimed to review zebrafish models of startle disease, identifying the genes and mutations involved in known zebrafish motility mutants. I also aimed to identify new mutations in the remaining alleles of the mutants *bandoneon* and *schlaffi*, exploring the possibility that the latter mutant harbours a mutation in *slc6a5* encoding GlyT2. I also aimed to explore new approaches for disrupting GlyR function in zebrafish, comparing morpholino knockdown to a new dominant-negative approach using a GlyR α4a subunit harbouring a dominant hyperekplexia mutation.

2. MATERIALS AND METHODS

2.1. DNA methodology

2.1.1. Materials

Ethanol 96% (cat #: 16368): Sigma-Aldrich Company Ltd (Dorset, UK)

QIAamp DNA Mini kit (cat #: 51304) containing QIAamp Mini spin column, buffer AL, buffer ATL, buffer AW1, buffer AW2, proteinase K and buffer EB: QIAGEN Ltd (Crawley, UK)

QiaQuick Gel Extraction kit (cat #: 28706) containing QiaQuick spin column, Buffer QG, Buffer PE and Buffer EB: QIAGEN Ltd (Crawley, UK)

SuperScript® III First-Strand Synthesis SuperMix (cat #: 18080-400) containing Oligo-dT, random hexamers, 2X First-strand reaction mix (10mM $MgCl_2$ and 1mM each dNTP), annealing buffer, SuperScript® III/RNase OUT™ enzyme mix: Invitrogen Ltd (Paisley, UK)

Advantage® 2 PCR kit (cat #: 639206), containing Advantage® 2 Polymerase mix, 10× Advantage® 2 PCR Buffer, 10 mM ultrapure dNTP mix: Clontech (Mountain View, USA)

Accuprime *Pfx* DNA polymerase (cat #: 12344), Invitrogen Ltd (Paisley, UK)

Betaine (cat #: B0300), Sigma-Aldrich Company Ltd (Dorset, UK)

UltraPure™ Agarose (cat #: 16500100), Invitrogen Ltd (Paisley, UK)

Thermo Hybaid PCR Express thermal cycler: Thermo Scientific (Loughborough, UK)

SYBRSafe® DNA gel stain (cat #: S33102): Invitrogen Ltd (Paisley, UK)

1kb plus DNA ladder (cat #: 10787-018): Invitrogen Ltd (Paisley, UK)

InGenius BioImaging system: Syngene (Cambridge, UK)

Safe Imager Transilluminator: Invitrogen (Paisley, UK)

Ligation Buffer (66 mM Tris-HCl, 5 mM $MgCl_2$, 1 mm dithiothreitol, 1 mM ATP, pH 7.5)

T4 DNA ligase (cat #: 10481220001): Roche Diagnostic Ltd (West Sussex, UK)

TOPO® TA cloning® kit for sequencing (cat #: K4575-02) containing pCR®4-TOPO® vector, salt solution and dNTPs: Invitrogen Ltd (Paisley, UK)

QuikChange II site-directed mutagenesis kit (cat #: 200523) containing *PfuUltra* High-fidelity DNA polymerase, 10X reaction buffer, *Dpn*I restriction enzyme and dNTP mix: Agilent (Cheshire, UK)

Ultrapure 25:24:1 (v/v) phenol:chloroform:isoamyl alcohol (PCI) mix (cat #: 15593): Invitrogen Ltd (Paisley, UK)

Acetic acid (cat #: A6283): Sigma-Aldrich Company Ltd (Dorset, UK)

Bromophenol blue (cat #: B0126): Sigma-Aldrich Company Ltd (Dorset, UK)

Ethylenediaminetetraacetic acid disodium salt (EDTA) (cat #: EDS): Sigma-Aldrich Company Ltd (Dorset, UK)

Sodium acetate anhydrous (cat #: S2889): Sigma-Aldrich Company Ltd (Dorset, UK)

Sucrose (cat #: S7903) Sigma-Aldrich Company Ltd (Dorset, UK)

Tris(hydroxymethyl)methylamine (Tris-base) (cat #: T1503): Sigma-Aldrich Company Ltd (Dorset, UK)

Restriction enzymes (e.g. *Dpn*I, *Eco*RI): New England Biolabs (Hitchin, UK)

Morpholinos: GeneTools LLC (http://www.gene-tools.com/), Oregon, USA.

Primers: Eurofins MWG Operon (http://www.eurofinsdna.com/), Ebersburg, Germany.

2.1.2. Preparation of genomic DNA

Tissue (up to 25mg) was placed into a 1.5 ml Eppendorf tube and lysed by adding 180 µl buffer ATL and 20 µl of proteinase K then incubated at 56°C with vigorous shaking for 3 hours or until all of the tissue was lysed completely. 200 µl Buffer AL was added to the lysed tissue solution and mixed by vortexing before incubating at 70°C for 10 minutes. 200 µl ethanol was added and the solution was mixed by vortexing. The DNA was bound to a QIAamp Mini spin column, then washed by adding 500 µl Buffer AW1 and centrifuging at 5,900 g for 1 min. Filtrate was discarded and the column was washed again by adding 500 µl Buffer AW2 and centrifuging at 16000 g for 3 min. Filtrate was discarded and the column was centrifuged again at 16000 g for 2 minutes to ensure that all ethanol was removed from the column. To elute the bound DNA, 100 µl Buffer EB was applied to the column and incubated at room temperature for 1 minute before centrifuging at 5,900 g for 1 minute. The spin column was discarded and the eluate containing the DNA was stored at -20°C.

2.1.3. Preparation of first-strand cDNA

In many instances, existing plasmid constructs containing DNA with the sequence of interest could be used as a template for the polymerase chain reaction. However, if an existing plasmid was not available, first strand cDNA was used which was synthesised from poly(A)+ RNA using SuperScript® III First-Strand Synthesis SuperMix. 1-5 µg

Poly(A)+ RNA, 1 μl primer (50 μM Oligo-dT / 2 μm gene specific primer / 50 ng/μl random hexamers) and 1 μl annealing buffer were added to a 0.2 ml PCR reaction tube and made up to a total volume of 10 μl with RNase/DNase-free water. The tube was mixed and incubated at 65°C for 5 min to allow the primers to anneal, then placed on ice. 10 μl 2× first-strand reaction mix and 2 μl SuperScript® III/RNase OUT™ enzyme mix were then added to the RNA/dNTP/primer mix. The final mixture was vortexed, centrifuged briefly then incubated at 50°C for 50 min (preceeded by an additional 10 min at 25°C if random hexamers used). To terminate the cDNA extension reaction, the tube was incubated at 85°C for 5 min then chilled on ice. The cDNA was then stored at -20°C until required.

2.1.4. Amplification using the polymerase chain reaction (PCR)

The Polymerase Chain Reaction (PCR) was used to amplify exons and flanking sequences for selected genes of interest from purified genomic DNA or full-length cDNAs from first-strand cDNA. Forward and reverse primers each of 20 - 30 bp in length were designed to cover the start and end of the region(s) to be amplified. All primer stocks were diluted to 10 pmol/μl and sequences of all primers used in this project can be found in the Appendix. Two PCR enzymes were used in this project: Advantage 2 Polymerase, a mixture of proofreading enzyme and the hot-start enzyme TITANIUM™ *Taq* polymerase. Reagents were added to PCR reaction tubes in the volumes shown in table 2.1.

Reagent	**Vol. per reaction**
Genomic DNA or cDNA	1 μl
10× Advantage 2 PCR Buffer	2.5 μl
10 mM ultrapure dNTP mix	0.5 μl
H_2O	19.75 μl
Forward and reverse primers (10 pmol/μl)	1 μl
Advantage 2 polymerase mix	0.25 μl
Total:	25 μl

Table 2.1. PCR reagents and volumes used. Reagent mix used in PCR amplifications from genomic DNA or cDNA.

5 µl betaine solution was added in addition to the reagents in table 2.1 if GC-rich regions of DNA were being amplified. Once all reagents had been added, the reaction tubes were centrifuged briefly and loaded into a Thermo Hybaid PCR Express thermal cycler. The temperatures and times of the stages of the PCR protocol were programmed in to the thermocycler and it was set to run. A general protocol of PCR cycling and descriptions of the stages are outlined in table 2.2.

Stage	Temp.	Time	Description
Denaturation	94°C	1 min	The dsDNA is heated to a high temperature at which the strands melt apart or denature allowing primers and deoxyribonucleotides (dNTPs) to bind to complementary base pairs.
Annealing	55-65°C	1 min	The primers anneal specifically to complementary regions on the ssDNA. The temperature depends on the TM of the primers. Temperature can be raised to reduce non-specific binding.
Extension	68-72°C	1 min per 0.5kb	The polymerase adds the free dNTPs to complementary base pairs on ssDNA strands.

Table 2.2. The stages of a polymerase chain reaction. The temperatures and times are a general guide and were optimised depending on the properties of the DNA being amplified and of the primers being used. These three stages make up one full cycle of PCR. The reaction requires a number of cycles, which is also dependant on the parent DNA, primers and desired amplicon, 25-40 cycles is generally acceptable.

2.1.5. Agarose gel electrophoresis

Gel electrophoresis is a method used to separate a mixed population of DNA fragments by length, to estimate the size of DNA fragments. An electric field moves negatively charged DNA molecules through an agarose matrix, with shorter molecules migrating farther and faster than longer ones through the pores of the gel. Solutions required: Tris-acetate EDTA (TAE) buffer - 40 mM Tris-base, 20 mM acetic acid, 1 mM EDTA, pH 8.0. 6× Loading buffer - 0.25% (w/v) bromophenol blue, 40% (w/v) sucrose in

distilled water. DNA fragments were visualized and separated using agarose gels containing SYBR Safe™ DNA gel stain. SYBR Safe™ DNA gel stain provides sensitive DNA and RNA detection with substantially reduced mutagenicity, making it safer to use than ethidium bromide. 2% (w/v) agarose gels were prepared by mixing 2 g of agarose powder into 100 ml of 1× TAE buffer and microwaving to dissolve the agar. After cooling the solution, 5 µl of SYBR Safe™ DNA Gel Stain was added to enable visualisation of DNA fragments after electrophoresis. The solution was poured in a casting tray containing a sample comb and allowed to solidify at room temperature for 30 min. The gel was placed in the electrophoresis chamber and immersed in 1× TAE buffer and the comb was removed. Samples were mixed with 3 µl of 6× loading buffer then loaded into the wells. 1 µg of 1 kb DNA ladder was used as a molecular weight marker. Samples were loaded in the wells and electrophoresed at a constant voltage of 100 V for 45 min. The DNA was visualised after this time under UV light using an InGenius BioImaging system or Safe Imager Transilluminator.

2.1.6. Purification of DNA from agarose gels

DNA bands were cut out from the gels using a clean sharp scalpel. Gel extraction was then carried out using the QiaQuick Gel Extraction kit. Gel slices were weighed and three gel volumes of buffer QG were added. The slices were incubated in buffer QG for 10 min at 50°C. One gel volume of isopropanol was then added and mixed. DNA was bound to a QiaQuick spin column matrix and washed with 0.75 ml of buffer PE. DNA was finally eluted in 20 µl of elution buffer (EB). To check the extraction, 5 µl of the eluted DNA was mixed with 3 µl of loading buffer and run on a 2% (w/v) agarose gel.

2.1.7. Restriction enzyme digestion

In order to prepare PCR inserts and plasmid vectors for ligation, DNA was digested using specific restriction endonucleases (Fig 2.1 stage 1). For PCR products, 15 µl of gel purified DNA was mixed with 2 µl 10 × enzyme buffer, 1 µl of each restriction enzyme and 2 µl dH_2O to a final volume of 20 µl. For vector DNAs, 2 µl plasmid DNA (at 0.5 µg/µl) was mixed with 2 µl 10 × enzyme buffer, 1 µl of each restriction enzyme (10

U/µl) and 15 µl dH_2O to a final volume of 20 µl. Mixtures were incubated at a specified temperature for a particular length of time depending on the enzymes, generally 37°C for 1-2 hours for most enzymes used. Restrictions enzymes for these digests were selected to give compatible ends for directional cloning of inserts into the plasmid vectors.

2.1.8. Phenol-Chloroform-Isoamyl alcohol extraction

This procedure was carried out to inactivate and remove restriction enzymes and buffer components before proceeding to the ligation reaction. The volume of the restriction enzyme digestion was increased to 100 µl with dH_2O. In a fume hood, an equal volume of Ultrapure Phenol/Chloroform/Isoamyl alcohol (PCI) mix was added to the DNA and the tube contents were mixed by tapping until an emulsion formed. The tube was then subjected to centrifugation at 16,000 g for 15 min. The upper aqueous phase was transferred to a clean 1.5 ml microcentrifuge tube containing 250 µl of 96% (v/v) ethanol, 10 µl 3 M sodium acetate (pH 4.8) and 1 µl of glycogen (1 µg). The contents were mixed thoroughly and the tube was incubated on dry ice for 15 min. The DNA was recovered by centrifugation at 16,000 g for 15 min and the supernatant was discarded. The pellet was washed with 250 µl 80% (v/v) ethanol and centrifuged at 16,000 g for 1 min. The supernatant was removed, the tube was subjected to a further centrifugation step at 16,000 g and any residual ethanol was removed. The pellet was air dried for 5 min at room temperature, dissolved in an appropriate volume of elution buffer EB and incubated at 37℃ for 10 min with shaking. When purifying vector DNA for subsequent ligation, the pellet was resuspended in 50 µl of buffer EB for a final vector concentration of 50 ng/µl. When purifying DNA inserts or PCR products, the pellet was resuspended in 10 µl of buffer EB.

2.1.9. Ligation of purified DNA inserts into a plasmid vector

After carrying out a restriction enzyme digestion and phenol-chloroform extraction on both the DNA insert and the plasmid vector, 7 µl DNA insert was added to 1 µl vector, 1 µl 10× ligation buffer and 1 µl bacteriophage T4 DNA ligase (1 U/µl) in a 1.5 ml

Eppendorf tube. Tubes were centrifuged briefly to mix and incubated at 4°C overnight to allow DNA and vector to ligate (Fig 2.1 stage 2).

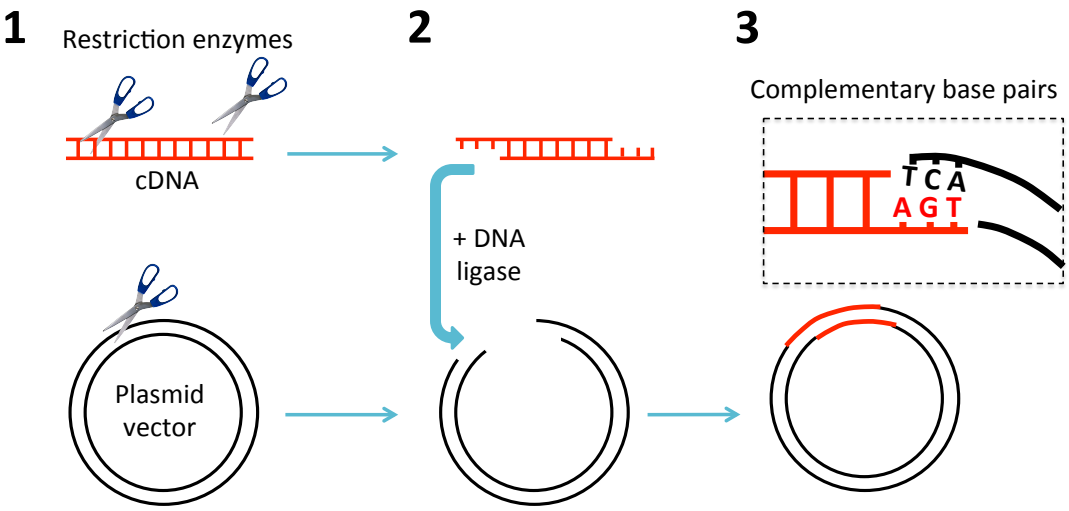

Figure 2.1. The stages of DNA ligation into a plasmid vector. 1. Restriction enzymes digest the cDNA and the plasmid vector DNA at specific sites leaving 'sticky ends' where each end has a number of unpaired nucleotides. After removal of restriction enzymes by PCI extraction, the cut cDNA and plasmid vector sequences are joined by complementary base pairing catalysed by the enzyme DNA ligase.

2.1.10. TOPO cloning

TOPO cloning is a fast and efficient cloning method that allows PCR products to be cloned without the need for restriction digestion. TOPO cloning utilises the enzyme topoisomerase I which functions both as a restriction enzyme and as a ligase. *Vaccinia* virus topoisomerase I specifically recognises the pentameric sequence 5'-CCCTT-3' and forms a covalent bond with the phosphate group attached to the 3' thymidine. It cleaves one DNA strand enabling the DNA to unwind. The enzyme then religates the ends of the cleaved strand and releases itself from the DNA. To harness the activity of topoisomerase, TOPO vectors are provided linearised with topoisomerase I covalently bound to each 3' phosphate. This enables the vectors to readily ligate DNA sequences with compatible ends. The TOPO TA Cloning kit for Sequencing was used allowing for

cloning of products with an adenine to the 3' end of the PCR product, as produced when Advantage® 2 polymerase was used. PCR products were run out on an agarose gel and then excised and purified as previously described. 4.5 µl of purified DNA was mixed with 0.5 µl pCR4-TOPO cloning vector and 1 µl of salt solution (1.2 M NaCl, 0.06 M $MgCl_2$) and incubated on ice for 30 min. The DNA mix was then transformed into *E. coli* competent cells as described in section 2.2.3.

2.1.11. Site-directed mutagenesis

Mutations were introduced into GlyT2 and GlyR expression constructs using the QuikChange site-directed mutagenesis kit. This rapid procedure uses a methylated supercoiled double-stranded DNA template and two synthetic oligonucleotide primers harbouring the desired mutation. The two primers, each complementary to opposite strands of the vector, are extended during temperature cycling by *PfuUltra* high-fidelity DNA polymerase. Subsequently, the PCR reaction is digested using the restriction enzyme *Dpn*I, which recognises and cuts methylated DNA from *E. coli*. Since only the non-mutated template DNA is methylated, the new PCR products containing the desired mutation are resistant to digestion and can be transformed into competent *E. coli*, which re-circularises the plasmid (Fig 2.2). Mutations were introduced into pRcCMV-hGlyT2, pRK5-hGlyR β or pCS2-zGlyR α4 plasmids using the QuikChange site-directed mutagenesis kit and specific primer pairs for each mutation. For each mutagenesis reaction, 5 µl of 10X reaction buffer, 50ng dsDNA template, 1 µl dNTP mix and 125 ng of each primer were added to a 0.2 ml microcentrifuge tube and the volume made up to a total of 50 µl with ddH_2O. After mixing, the tube was briefly centrifuged and placed in a Thermo Hybaid PCR Express thermal cycler. The PCR reaction was carried out using the following conditions: an initial denaturation at 95℃ for 1 min was followed by 18 cycles of denaturing at 95℃ (1 min), annealing at 60℃ (1 min) and extension at 72℃ (16 min).

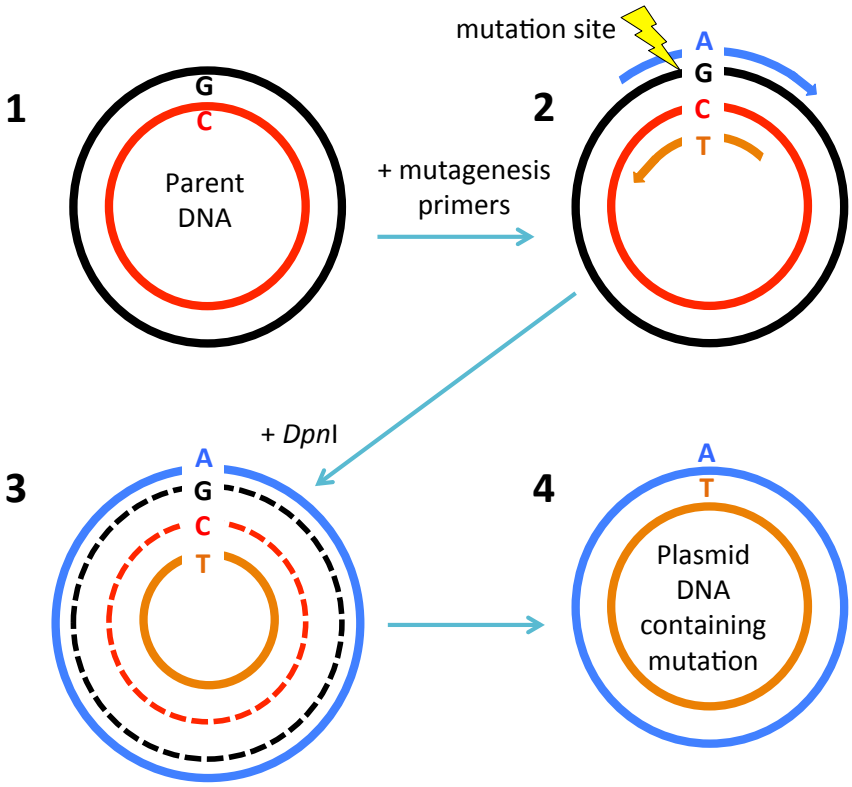

Figure 2.2. The stages of site-directed mutagenesis. 1. Parent DNA template is denatured. **2**. Mutagenesis primers containing the desired mutation anneal to the strands of the parent DNA and the new DNA extends. **3**. Restriction enzyme *Dpn*I digests the methylated parent DNA, leaving the new DNA containing the mutation intact. **4**. The mutant DNA is then transformed into competent *E. coli* cells which repairs nicks in the DNA to form circular plasmids.

PCR products were digested with *Dpn*I in order to eliminate methylated template DNA and select for newly synthesised DNA containing the desired mutation. 1 µl *Dpn*I (20 U/µl) was added to the 50 µl mutagenesis PCR product, mixed thoroughly and incubated for 1 hour at 37℃. Part of the *Dpn*I digested PCR product (10 µl) was then loaded on a 1% (w/v) agarose gel and subjected to electrophoresis at 90 V for 30 min. Detection of a linear DNA fragment with a size consistent with vector plus insert indicated the presence of mutated DNA, since all non-mutated DNA was removed by digestion with *Dpn*I. Finally, 10 µl from each reaction were transformed into *E. coli* competent cells and miniprep DNAs were made for analysis (see section 2.2 for methods). The full coding region of all constructs was sequenced to confirm the

incorporation of the desired mutation and to ensure that no unwanted changes had been incorporated via PCR errors.

2.1.12. Design of morpholinos for zebrafish gene knockdown

Morpholinos (MOs) are short (25mer) nuclease-resistant oligonucleotides that block access of other molecules to specific RNA sequences, which can either prevent cells from making a target protein (translation blocking MOs) or modify the splicing of pre-mRNAs (splice MOs). Microinjection of MOs into zebrafish embryos (Nasevicius & Ekker 2000; Bill *et al* 2009) is now an established tool for studying gene function. However, key controls are required, since common non-specific effects of MOs at higher amounts (>4-10 ng per injection) include a general delay in development or widespread cell death (Bill *et al* 2009). We therefore designed and synthesized both translation-blocking and splice-blocking MOs specific for GlyR α4a by aligning exon 7 sequences for zebrafish GlyR genes (Fig. 2.3). The morpholino sequences used were: *Glra4a*SMO1 5'-acctagaagagcacaaagagtttca-3', *Glra4a*SMO2 5'-acaggaactcattttatgttacctt-3', *Glra4a*TBMO 5'-aaatccttatgacctgagggagcat-3'. We also determined the optimal MO amount required for induction of a specific phenotype post-injection, without inducing toxicity. RT-PCR was also used to assess the efficacy of splice morpholinos. As a final control, a dominant-negative mutation (R278Q) was introduced into a zebrafish GlyR α4a expression construct in pCS2 and this construct was also microinjected into zebrafish embryos.

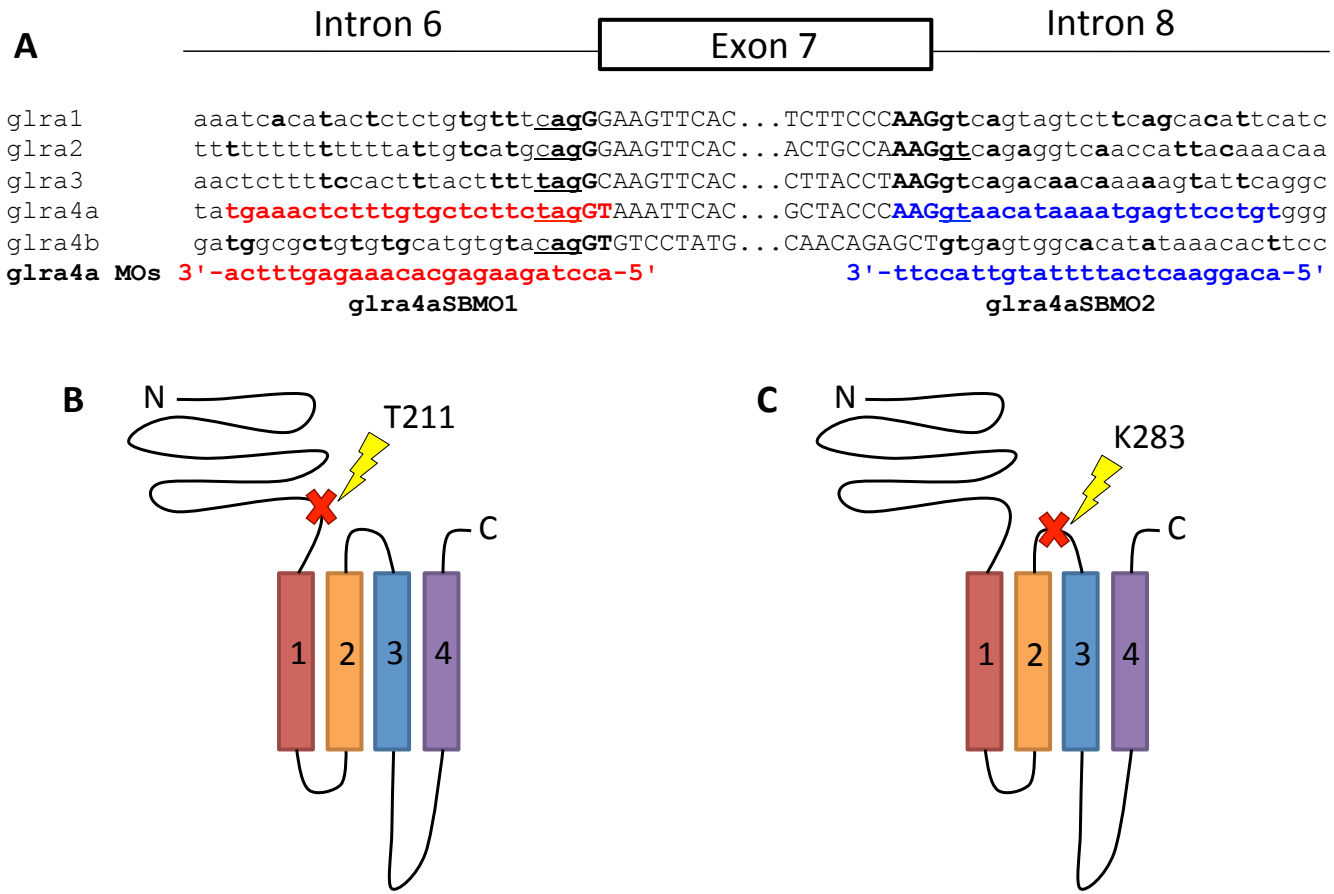

Figure 2.3. Morpholino oligonucleotides designed to block splicing of *glra4a*. SMO1 and SMO2 block the donor (ag) or acceptor (gt) sites for exon 7. This is predicted to result in intron retention, mis-splicing or exon skipping. In any of these events, the GlyR polypeptide would be truncated as shown in panels B/C – see chapter 6 for further details.

2.2. Bacterial methodology

2.2.1. Materials

OneShot® TOP10 chemically competent *E. coli* cells (genotype: F- *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ *lac*X74 *rec*A1 *ara*D139 Δ(*araleu*)7697 *gal*U *gal*K*rps*L (StrR) *end*A1 *nup*G) (cat #: C4040): Invtrogen Ltd (Paisley, UK)

Ampicillin (cat #: 11593-027): Invitrogen Ltd (Paisley, UK)

Kanamycin (cat #: 11815-024): Invitrogen Ltd (Paisley, UK)

Luria-Bertani (LB) medium (cat #: L3022) (1% (w/v) peptone, 0.5% (w/v) yeast extract, 0.5% (w/v) NaCl): Sigma-Aldrich (Dorset, UK)

Luria-Bertani (LB) Agar (cat #: L2897) (LB medium supplemented with 1.5% (w/v) agar): Sigma-Aldrich (Dorset, UK)

$CaCl_2$ (cat #: C1016): Sigma-Aldrich (Dorset, UK)

Glycerol (cat #: G5150): Sigma-Aldrich (Dorset, UK)

$MgCl_2$ (cat #: M8266): Sigma-Aldrich (Dorset, UK)

QIAprep Spin Miniprep kits (cat #: 27106) containing buffer P1 (resuspension buffer - 50 mM Tris-Cl (pH 8.0), 10 mM EDTA, 100 µg/ml RNase A), buffer P2 (lysis buffer - 200 mM NaOH, 1% (w/v) SDS), buffer N3 (neutralisation buffer - 3 M potassium acetate (pH 5.5)), Buffer PE (wash buffer - 1 M NaCl, 50 mM MOPS (pH 7.0), 15% (v/v) isopropanol) and Elution Buffer (EB – 10 mM Tris-HCl (pH 8.5)): Qiagen Ltd (Crawley, UK)

HiSpeed Plasmid Maxi kits (cat #: 12663) containing Qiafilter Cartridge, HiSpeed Maxi Tip, Qiaprecipitator, buffers P1, P2, P3, EB, QBT (equilibration buffer - 750 mM NaCl, 50 mM MOPS (pH 7.0), 15% (v/v) isopropanol, 0.15% (v/v) Triton X-100), buffer QC (wash buffer), buffer QF (elution buffer - 1.25 M NaCl, 50 mM Tris-HCl (pH 8.5), 15% (v/v) isopropanol): Qiagen Ltd (Crawley, UK)

*Eco*RI, F*au*I and *Hpy*CH4V restriction enzyme: New England Biolabs (Hitchin, UK)

10× *Eco*RI buffer: New England Biolabs (Hitchin, UK)

AccuPrime™ *Pfx* Super mix : Invitrogen (Paisley, UK)

LightScanner® Master Mix (cat #: HRLS-ASY): BioFire Diagnostics Inc (Salt Lake City, USA)

LCGreen plus melting dye (cat #: BCHM-ASY): BioFire Diagnostics Inc (Salt Lake City, USA)

HotShot master mix (cat #: HS002): Clent Life Science Ltd (Stourbridge, UK)

LightScanner 96 High Resolution Melting instrument: BioFire Diagnostics Inc (Salt Lake City, USA)

Nanodrop UV spectrophotometer

Sequencher 5.0 software: GeneCodes Corporation (Ann Arbor, USA)

2.2.2. Bacterial culture

LB medium and LB agar medium were used for routine bacteria culture. For plasmid selection, LB medium or agar cultures were supplemented with antibiotics at a final concentration of 100 µg/ml (ampicillin) or 50 µg/ml (kanamycin).

2.2.3. Preparation of chemically competent Escherichia coli cells

Solutions required: 80 mM $CaCl_2$/50 mM $MgCl_2$, 0.1 mM $CaCl_2$, 50% (v/v) glycerol. Competent *E. coli* cells were prepared using a modified $CaCl_2$ protocol as described by Cohen *et al* (1972). TOP10 *E. coli* cells were used to produce competent cells. Aseptic technique was used to streak TOP10 cells onto LB agar plates without antibiotics, which were then incubated at 37°C overnight. Single colonies were picked and inoculated into 3 ml LB broth then incubated in a shaking incubator at 37°C with shaking overnight. The cultures were transferred into 500 ml conical flasks containing 200 ml of pre-warmed LB broth and then incubated in a shaking incubator at 37°C. The optical density (OD) of the cultures at 600 nm were measured using a Eppendorf Biophotometer after 2 hours and then regular intervals following this until the OD_{600} of the cultures reached 0.95. As soon as this OD was reached, the cultures were transferred into 50 ml falcon tubes that had been incubated on ice and centrifuged at 2,700 g at 4°C for 5 min using a bench top centrifuge. Supernatant was poured off and

the pellets were put on ice. Pellets were resuspended in 10 ml of 80mM $CaCl_2$/50mM $MgCl_2$ solution, and pooled into a single falcon tube. The cells were incubated on ice for 10 min before centrifuging again at 1,500 g at 4°C for 3 min. This process was repeated twice, in order to wash the cells in the 80 mM $CaCl_2$/50 mM $MgCl_2$ solution a total of three times. The cells were then resuspended in 5.5 ml of ice-cold 0.1 mM $CaCl_2$ solution and an equal volume of ice-cold 50% (v/v) glycerol was added. The cells were then dispensed into 550 µl aliquots in pre-chilled microcentrifuge tubes and frozen immediately in liquid nitrogen. Competent cells were stored at -80°C until required.

2.2.4. Transformation into competent Escherichia coli cells

Bacteria were transformed using a protocol based on work of Mandel and Higa (1970) and Cohen *et al* (1972) who showed that bacteria treated with ice-cold $CaCl_2$ and then subjected to a brief heat-shock to take up plasmid DNA. Vials of chemically competent *E. coli* cells were gently thawed on ice. 1-5 µl of the DNA to be transformed was added to 100 µl cells and incubated on ice for 30 min. Cells were heat shocked by placing in a water bath at 42°C for 45 sec and then transferred immediately to ice for 2 min. 250 µl of room temperature LB broth was then added and the bacteria were incubated at 37°C for 1 hr in a shaking incubator. During this step, the cells were able to express the antibiotic resistance gene encoded by the plasmid. Finally, bacteria were then spread on LB agar plates containing ampicillin or kanamycin using sterile technique and the plates were inverted and incubated at 37°C overnight.

2.2.5. Small-scale preparation of plasmid DNA (minipreps)

DNA minipreps were made using the Qiagen Spin Miniprep kit. This is based on the alkaline lysis of bacteria followed by the absorption of DNA onto a silica membrane and elution of DNA. Briefly, single colonies from the transformed agar plates were picked using a pipette tip and used to inoculate 2 ml of LB medium supplemented with the appropriate antibiotic in a 15 ml falcon tube. After incubation of the culture at 37°C overnight with vigorous shaking, bacterial cells were pelleted by centrifugation at

16,000 g for 1 min. The supernatant was discarded and the pellet was resuspended in 250 µl pre-chilled buffer P1 containing lysozyme and RNase. 250ul of Buffer P2 was then added and mixed by inversion. 350ul of neutralization buffer N3 was finally added and mixed by inversion. The microcentrifuge tubes were then centrifuged at 16000 g for 10 minutes. The supernatant was then transferred to a mini spin column and centrifuged at 16000 g for 1 minute in order to bind the DNA. The column was then washed with 750ul of wash buffer PE in order to remove any cellular debris and two centrifugation steps were performed at 16000g for 1 minute each to remove all wash buffer and ensure that the column was dry. The column was placed in a fresh microcentrifuge tube and 50 µl of buffer EB was added to the matrix. To elute the DNA, the tube containing the column was centrifuged at 16,000 g for 1 min. The yield was determined using a Nanodrop UV spectrophotometer and typically varied from 0.2 to 0.8 µg/µl. Miniprep DNAs were then stored at -20°C until required. Miniprep DNAs were determined to contain correctly sized DNA inserts by digestion with restriction endonucleases, such as *Eco*RI. 5 µl of purified plasmid DNA was mixed with 0.5 µl *Eco*RI, 1 µl 10× *Eco*RI buffer and 3.5 µl water and incubated at 37°C for 1 hour. Digested DNAs were then mixed with 3 µl loading buffer and run on a 2% (w/v) agarose gel and visualised as before.

2.2.6. Large-scale preparation of plasmid DNA (maxipreps)

Maxipreps of plasmid DNA were carried out according to the HiSpeed Plasmid Purification procedure, which is based on a modified alkaline lysis method followed by the binding of plasmid DNA to a resin in low salt and pH conditions. A single bacterial colony was picked from the agar plate and resuspended in 2 ml of LB medium supplemented with the appropriate antibiotic in a 15 ml falcon tube. The tube was incubated for approximately 8 h at 37°C in a shaking incubator. A 200 µl volume of the bacteria culture was transferred into a sterile flask containing 200 ml of LB medium and the appropriate antibiotic. The culture was grown for 12-16 hours at 37°C in a shaking incubator. Bacterial cells were harvested by centrifugation at 4,000 g for 15 min at 4°C. After removing the supernatant, the pellet was resuspended in 10 ml pre-

chilled buffer P1. Then 10 ml of lysis buffer P2 was added and the suspension was mixed by inverting six times and incubated at room temperature for 5 min. Finally, 10 ml of pre-chilled buffer P3 was added to the lysate, the suspension was poured into the barrel of the Qiafilter Cartridge and incubated at room temperature for 10 min. During this incubation step, a HiSpeed Maxi Tip was equilibrated by applying 10 ml buffer QBT, allowing the column to empty by gravity flow. After removing the cap from the Qiafilter outlet nozzle, the cell lysate was filtered into the previously equilibrated HiSpeed Tip and allowed to enter the resin by gravity flow. After washing the HiSpeed Tip with 60 ml Buffer QC, the DNA was eluted with 15 ml Buffer QF and the eluate was collected in a clean 50 ml falcon tube. DNA was precipitated by adding 10.5 ml isopropanol to the eluate and the mix was incubated at room temperature for 5 min. During this incubation, a Qiaprecipitator Maxi Module was attached to a 30 ml syringe. The eluate/isopropanol mixture was filtered through the Qiaprecipitator. DNA was washed by adding 2 ml 70% (v/v) ethanol to the syringe. After drying the membrane, the Qiaprecipitator was attached to a 5 ml syringe and 0.5 ml EB buffer was added to the syringe. The DNA was eluted into a fresh 1.5 ml microcentrifuge tube. The eluate was transferred into a syringe and the elution step was repeated in order to maximise the amount of DNA solubilised and recovered from the Qiaprecipitator. The DNA concentration was determined by spectrophotometry at 260 nm and varied between 0.5 and 4 µg/µl. Maxiprep DNAs were then stored at -20°C until required.

2.2.7. Detection of the congenital muscular dystonia type 2 (CMD2) mutation using RFLPs and LIghtscanner HRM analysis

For RFLP detection, standard PCRs were performed using 50 ng genomic DNA, *Pfx* supermix and the primers bGlyT2Ex4F1 (5'-ttgggcctctctgagctctc-3') and bGlyT2Ex5R1 (5'-ccccagcatctagtatagagcc-3') with 35 cycles of 94°C for 1 min, 60°C for 1 min, 68°C for 2 min. PCR products were gel purified using a QiaQuick gel extraction kit before restriction digestion with F*au*I or *Hpy*CH4V or Sanger DNA sequencing (http://www.dnaseq.co.uk/). For standard HRM analysis, PCR was performed using 25 ng genomic DNA, LightScanner Master Mix and primers bGlyT2Ex4F2 (5'-

gctctcacctcctactctctctttccaa-3') and bGlyT2Ex4R1 (5'-gcagggaagggctgaggc-3') with 45 cycles of 94°C for 30 sec and 70°C for 30 sec. For LunaProbe analysis, PCR was performed using HotShot master mix in the presence of LCGreen Plus Dye and the primers bGlyT2Ex4F1 (5'-ttgggcctctctgagctctc-3') and bGlyT2Ex4R2 (5'-gaacagggtgatgccagagatg-3') and unlabelled 3' blocked probe CMD2P (5'-actcaccttgcagggctgggat-3') with 55 cycles of 95°C for 30 sec, 64°C for 30 sec and 72°C for 30 sec. The primers were run asymmetrically at a ratio of 5:1 in favour of the forward primer. The resulting amplified products were then melted from either 85°C to 97°C (standard HRM) or 50°C to 95°C (LunaProbe analysis) on the LightScanner 96 High Resolution Melting instrument. Analysis was performed using Idaho Technology Call-IT 2.0 Software.

2.2.8. DNA sequencing and analysis

Sequencing of DNA samples was carried out by Sequencing Services, School of Life Science, University of Dundee, Scotland (http://www.dnaseq.co.uk/). DNA sequencing was performed using Applied Biosystems Big-Dye Version 3.1 chemistry on an Applied Biosystems model 3730 automated capillary DNA sequencer (Life Technologies, Paisley, UK). Template DNA was supplied at a concentration of 200-300 ng in a 15 µl volume per reaction (3.2 pmol). Sequencing primers (e.g. M13F or M13R) were either provided by the service, or sent at 3.2 pmol per reaction. DNA templates and primers were diluted in dH_2O. DNA sequences were analysed using the software Sequencher 5.0 by alignment with reference sequences downloaded from NCBI (http://www.ncbi.nlm.nih.gov/) or UCSC databases (http://genome.ucsc.edu/). For mutation identification, single nucleotide variants (SNVs) were noted and the effect of these changes on the encoded protein were examined using bioinformatics programmes such as SIFT (http://sift.jcvi.org/) and Polyphen-2 (http://genetics.bwh.harvard.edu/pph2/index.shtml) prior to molecular modelling.

2.3. Functional assays

2.3.1. Materials

Protein assay kit (cat #: 500-0001): Bio-Rad Laboratories UK Ltd (Hemel Hempstead, UK)

Minimal essential medium (MEM): Invitrogen Ltd (Paisley, UK)

Foetal bovine serum (FBS): PAA Laboratories Ltd (Yeovil, UK)

Lipofectamine LTX transfection reagent kit : Invitrogen Ltd (Paisley, UK)

Glycine (cat #: 50046): Sigma-Aldrich Company Ltd (Dorset, UK)

Penicillin-streptomycin solution (cat #: P4333): Sigma-Aldrich Company Ltd (Dorset, UK)

Phosphate buffered saline (PBS) tablets (cat #: P4417): Sigma-Aldrich Company Ltd (Dorset, UK)

Poly-L-lysine (cat #: P4707): Sigma-Aldrich Company Ltd (Dorset, UK)

Human Embryonic Kidney (HEK) 293 cells were obtained from the European Collection of Cell Cultures (Salisbury, UK)

2.3.2. [^{3}H]glycine uptake assays

Solutions required: Krebs buffer - 118 mM NaCl, 1mM NH_2PO4, 26 mM $NaHCO_3$, 1.5 mM $MgSO_4$ mM, 5 mM KCl, 1.3 mM $CaCl_2$, 20 mM glucose. HEK293 cells were cultured in MEM medium supplemented with penicillin-streptomycin solution and plated on poly-L-lysine coated 24-well plates. When 40% confluent, HEK293 cells were transfected with the pRcCMV-GlyT2 plasmids at 1 µg/µl per well with Lipofectamine LTX reagent and incubated at 37°C. 24 hours after transfection, cells were washed twice with 1 ml/well warm Krebs buffer pre-equilibrated with 5% CO_2-95% air. After 2 min, cells were incubated with 1 ml/well [^{3}H]glycine at a final concentration of 300 µmol (0.1 µCi/ml) for 5 min. Cells were rinsed twice with ice-cold Krebs buffer pre-equilibrated with 5% CO_2-95% air, then they were digested with 1 ml/well of 0.1M NaOH for 2 hours. 750 µl of each sample was pipetted into tubes and 3.5 ml scintillation fluid was added to each tube. The radioactivity incorporated was

measured by scintillation counting. 50 µl of each sample was used for determining the protein concentration using the Bradford reagent (BioRad). [^{3}H]glycine uptake was calculated as nmol/min/mg protein and expressed as percentage of that in control cells transfected with the empty vector. All statistical comparison used an unpaired Student's t-test.

2.4. Molecular Modelling

2.4.1. Template detection (fold recognition)

Homologues to the protein to be modelled with known structures were found using HHSearch and GenTHREADER. HHSearch is part of the HHPred online web server (http://toolkit.tuebingen.mpg.de/hhpred/) that uses pair-wise comparison of Hidden Markov Models (HMM) profiles for remote protein homology detection and structure prediction (Söding *et al* 2005). The target sequence was entered into the online form and all parameters were set to default, where the most recent HMM database is used, multiple sequence alignment is calculated with HHblits with a maximum of 3 iterations, secondary structure confidence is scored and a local alignment mode is used. Results from the HHPred online server were returned listing the top 10 closest related structures. Each structure was examined for percentage sequence identity and biological function and the best structure or structures were selected as templates.

Additional template detection was performed with GenTHREADER (http://bioinf.cs.ucl.ac.uk/psipred/). This server makes use of a different method for predicting protein structure, namely protein threading (Jones 1999; Lobley *et al* 2009). Threading is generally used for proteins for which a structural homologue that could be used as a template structure cannot be easily detected by sequence alignment methods (Jones & Thornton 1994). Since the number of different known native folds is quite small (around 1300 according to the CATH Protein Structure Classification database (http://www.cathdb.info) (Orengo *et al* 1997), and very few are added each year), there is a high possibility that any given protein with an unsolved structure can be classified as having a fold similar to another protein that already has a solved

structure that is available in the PDB. GenTHREADER searches for structures that have a similar fold by 'threading' each amino acid in the target sequence to a position in the template structure, while evaluating the fit using a complex scoring function (Jones 1999, McGuffin & Jones 2003). The resulting sequence alignment could then be used to create a model, but typically it is better to use improved sequence alignment method for this task. GenTHREADER was used to identify structures of distantly related proteins to the target sequence that might not have been found using HHPred.

2.4.2. Creating profile alignments for the target and template sequences

Creating a good quality sequence alignment between a query protein and a distant homologue template protein is essential for its effectiveness for structure prediction. NCBI Protein BLAST (Basic Local Alignment Search Tool; http://blast.ncbi.nlm.nih.gov) (Altschul *et al* 1990) was used to search the non-redundant protein sequence database using the position-specific iterated BLAST (PSI-BLAST) algorithm for proteins similar to that of the target sequence and the template sequence. The 50 most closely related sequences were selected and the sequences downloaded. Multiple sequence alignments were calculated for both the target sequence and the template sequence (each with the 50 most homologous protein sequences) using the MUSCLE (MUltiple Sequence Comparison by Log-Expectation) web server (http://www.ebi.ac.uk/Tools/muscle/index.html; Edgar 2004).

2.4.3. Creating profile-profile alignments of the target and template

Profile-profile alignments were performed in order to align the target sequence with that of the template more accurately and considering areas of conservation or evolutionary changes within the families. One single alignment was calculated by uploading each of the two profile alignments in 'Fasta' format into the T-COFFEE web server (Notredame *et al* 2000) using the 'Combine' option. The output was examined and all sequences apart from those of the target and template were removed, leaving just two aligned sequences.

2.4.4. Assessment of alignment and preparation of alignment file

The final alignment was checked visually in the protein visualisation program Chimera (Petterson *et al* 2004) considering the secondary structures of the template structure (actual structure) and of the target structure (predicted secondary structure). Secondary structure of the target sequence was predicted using the TMHMM web server, a server designed to predict transmembrane helices in proteins using HMM (http://www.cbs.dtu.dk/services/TMHMM/) (Krogh *et al* 2001). Residues that were not optimally aligned based on the actual and predicted secondary structures were manually realigned. Alignment files were edited in order to allow the model-building program to read them correctly. The total number of aligned residues and gaps were checked to be the same for each sequence and the file was saved with a '.pir' extension allowing the modelling program to read it. The 'pir' (or 'ali') format must provide specific information about the target and template sequences. The first line preceding each sequence describes the sequences themselves, annotating the protein names, chains and residues (Fig 2.4).

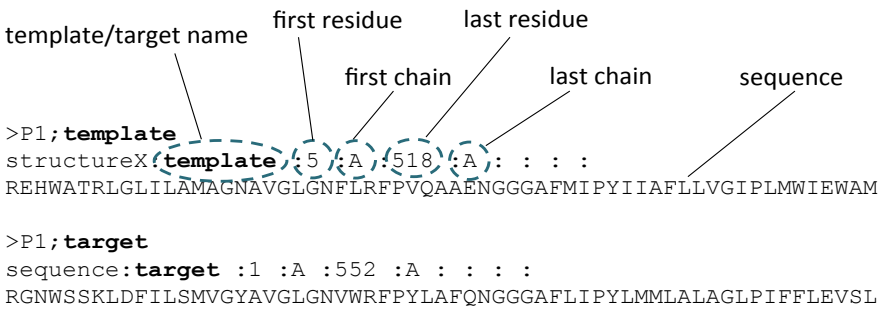

Figure 2.4. MODELLER PIR file format. An example alignment in the 'pir' format required for utilisation by MODELLER. The first line describes the sequence type, P1 in this case which is the annotation for protein. This is followed by a semicolon and the sequence identity. The second line contains information about sequence numbering and chain codes with each element of information separated by a colon. First, 'structureX' for the template or 'sequence' for the target. 'X' in 'structureX' is arbitrary, but if multiple templates are used it is replaced with a different letter for each template. Next, the sequence I.D., which can be the PDB reference I.D. if available. This is followed by the start residue, start chain I.D., end residue and end chain I.D. Finally there may be a description of the sequence or structure, if desired, for example, the full name of the protein or the experimental procedure used for the determination of the structure. This is written into the coordinate file of the final model.

2.4.5. Preparation of a Python script for the alignment to be read into the model-building program MODELLER

MODELLER (Sali & Blundell 1993) is a popular computer program designed for building homology models. This software builds homology models using an alignment of the target sequence to the template structure, the atomic coordinates of the template structure and a simple script to instruct the program (Sali & Blundell 1993; Eswar *et al* 2006). The program executes the model building by the satisfaction of spatial restraints, whereby a set of geometrical criteria help create a probability density function for each atom in the structure. MODELLER uses Python as a control language; therefore all input scripts must be Python scripts. A Python script was written to instruct MODELLER to create a specified number of models based on the alignment of the target sequence with the template structure. An example of the script used is demonstrated in Fig 2.5.

```
from modeller.automodel import *

log.verbose()
env = environ()

A  env.io.atom_files_directory = '/Users/victoriajames/Work/BBK/glyt2/modeller'
B  env.io.hetatm = True

C  a = automodel(env,
D                 alnfile  = 'glyt2_2A65.ali',
                  knowns   = ('2A65'),
                  sequence = 'GlyT2', assess_methods=(assess.DOPE, assess.normalized_dope))
E  a.starting_model= 1
   a.ending_model  = 50
   a.make()

   ok_models = filter(lambda x: x['failure'] is None, a.outputs)

F  key = 'DOPE score'
   ok_models.sort(lambda a,b: cmp(a[key], b[key]))

   m = ok_models[0]
   print "Top model: %s (DOPE score %.3f)" % (m['name'], m[key])
```

Figure 2.5. An example python script used to instruct MODELLER to build models based on an alignment. The script instructs MODELLER to build 50 models for GlyT2 based on the alignment of the sequence with $LeuT_{Aa}$ (PDB I.D.: 2A65). The script describes the location of the input files (**A**), heteroatom inclusion is 'true' to include Na^+ ions (**B**), the alignment file name ('glyt2_2a65.ali') and the names of the template structure ('2A65') and target sequence ('GlyT2') as specified in the alignment file (**C**), 'automodel' class is used which automates the steps in the model building (**D**), 50 models should be calculated (**E**), and finally the DOPE score for each model should be calculated and the model with the lowest DOPE score selected (**F**).

In brief, this script describes the location of the input files (Fig. 2.5a), whether heteroatoms should be included or not (Fig. 2.5b), the alignment file name and the names of the template structure and target sequence (Fig 2.5d), the number of models that it should calculate (Fig. 2.5e), and finally to calculate the discrete optimised potential energy (DOPE) score for each model (Fig 2.5f and section 2.4.6). The raw DOPE score is an atomic distance-dependant statistical potential derived from a sample of native protein structures (Shen & Sali 2006). The scoring function incorporates a probability density function of the atomic coordinates and bond lengths together with a probability density function of interactions derived from a sample of native structures. It is particularly useful in this method to rank models made in a run. The model with the lowest DOPE score was selected and reported to the log output file (Fig. 2.5f). All the files that MODELLER required to run the script, including the alignment file and the PDB atomic co-ordinate file for the template structure, were placed in the same directory with the script. MODELLER was instructed to run the specified script via the command line terminal. Models were calculated and the coordinate files were listed in the same directory as the script automatically.

2.4.6 Assessing the quality of the models by creating an energy profile using DOPE

The DOPE (Discrete Optimised Protein Energy) statistical potential (Shen & Sali 2006) is an atomic distance-dependent statistical potential that is derived from a sample of native structures in the protein data bank (PDB; www.pdb.org; Berman *et al* 2000). It is incorporated within MODELLER so that a DOPE score may be calculated following model building (Fig. 2.5e) for each model built. MODELLER was instructed to calculate the energy of the spatial arrangements of the atoms in each residue in the final model. The same energy profile was also calculated for the template structure. The energies for both the new model and the template structure were plotted against residue number using the graphical program Gnuplot (http://www.gnuplot.info/). Plots were then assessed to determine areas of the model that could be improved. Tall peaks or positive values representing high energy, in general, correspond to problematic parts of a model (Wiederstein & Sippl 2007). Overlapping peaks and troughs of a model

profile with the template profile corresponds to well aligned areas of the sequence. An example of an energy profile created for the X-ray crystal structure of the *E. coli* arginine:agmatine antiporter (PDB I.D. 3LRB) is shown in Fig 2.6.

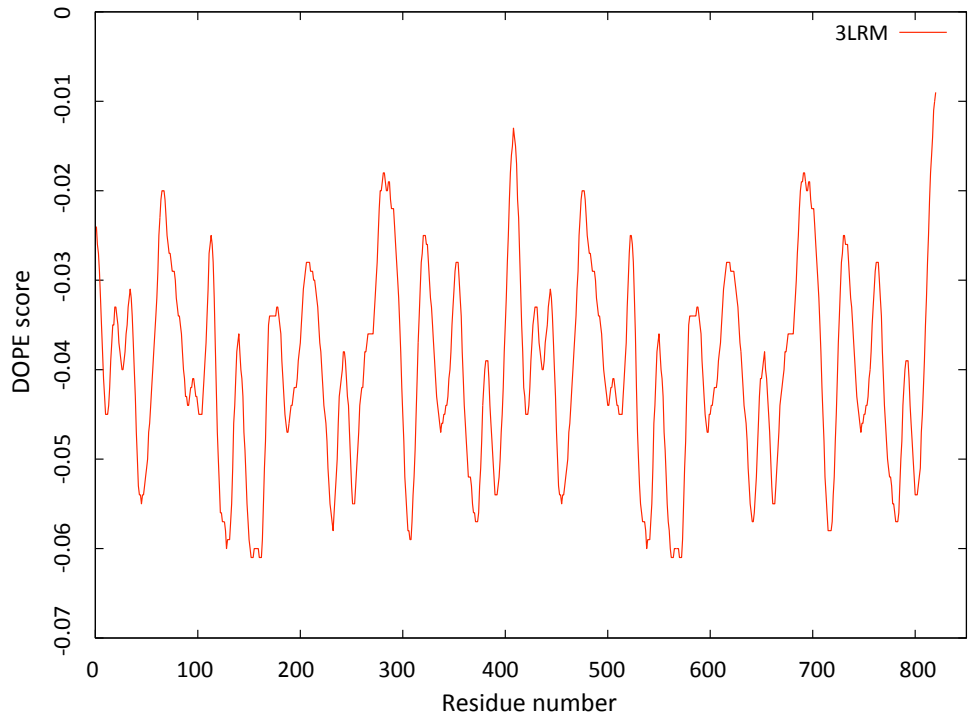

Figure 2.6. Energy profile for the X-ray crystal structure of the *E. coli* arginine:agmatine antiporter (PDB I.D. 3LRB). The DOPE energy for each atom in the structure is calculated then plotted in sequence. Peaks of high energy, particularly values that are close to or over 0, are considered unfavourable.

2.4.7. Calculation of a normalised DOPE Z-score for prediction of overall accuracy of a model

Computationally derived homology models typically contain significant errors particularly when there is a low sequence identity with the template (Eramian *et al* 2008). The overall accuracy of a model is important to realise if it is to be used to provide any biological insight. Methods by which the accuracy of a model can be assessed are still somewhat limited but there are ways that a model can be scored and assessed to determine its goodness. A normalized DOPE Z-score was calculated for the best model within MODELLER using a simple script. This score is appropriate for an

overall assessment of a models quality by normalising the raw DOPE scores of a model over its sequence length. This gives some assessment of a model quality by calculating a positive or negative Z-score, where positive scores are likely to represent poor models whereas negative scores, particularly below -1 are likely to represent a native-like structure.

2.4.8. Additional assessment of the final model using QMEAN

Additional evaluation of the homology model was performed using the QMEAN web server (http://swissmodel.expasy.org/qmean/cgi/index.cgi) (Benkert *et al* 2008, 2009). The .pdb file containing atomic coordinates for the model was uploaded on to the server and the 'QMEAN' scoring function option was selected. A QMEAN score between 0 and 1 is returned reflecting the predicted model reliability, with 1 being most reliable and 0 being the least reliable. The QMEAN Z-score is also returned which can be positive or negative, where high-resolution structures have a score of around 0, with scores becoming worse as they become more negative. The Z-score is evaluated by comparing it to the Z-scores calculated for native proteins that have high-resolution structures solved by X-ray crystallography (Benkert *et al* 2009, 2011). The sample protein Z-scores are plotted against protein size and the Z-score for the model is plotted on the same graph (Fig 2.7a). Other helpful results are also reported such as the QMEAN scores for specific terms that make up the overall score such as Cβ interaction, all atom interaction and torsion angles. A residue error plot is also generated, similar to the plot generated for the individual residue DOPE scores, which highlights areas of the model that have higher predicted error. In addition, a coordinate file is created containing this additional information allowing a 3D structure to be coloured according to predicted error (Fig. 2.7b).

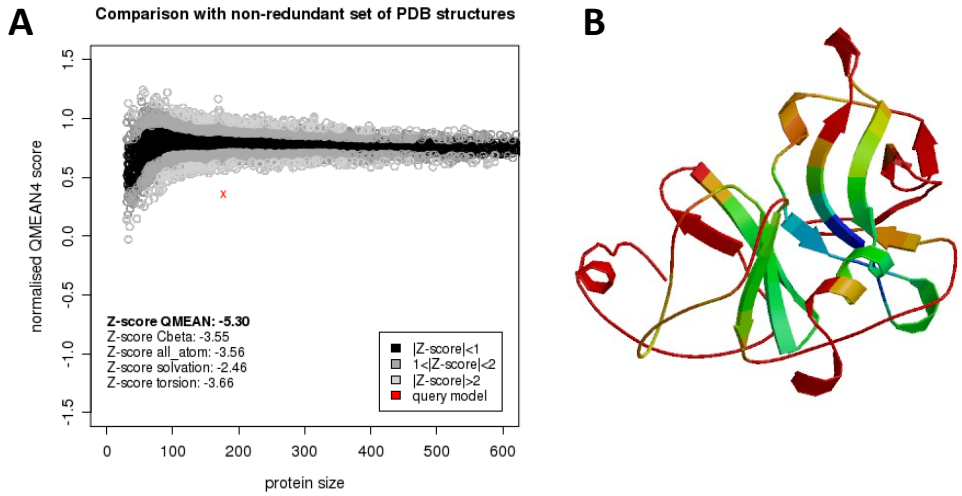

Figure 2.7. Example output from the QMEAN server. QMEAN quality estimation of the experimentally solved protein structure of the Dengue NS3 serine protease (PDB I.D. 1BEF) **A**: Plot of QMEAN Z-scores of native proteins with structures solved by X-ray crystallography against protein size (black, dark grey and light grey circles). The QMEAN Z-score for the query model is also plotted for comparison (red cross). **B**: The structure of the query model coloured by residue error, where red corresponds to high error through the colour spectrum to blue, which indicates low error.

2.4.9 Energy minimization of a structure in Chimera

Energy minimization was carried out on the model with the *minimize structure* command in Chimera, using the Molecular modelling toolkit (MMTK) method (Hinsen 2000) and the AMBER94 force field (Cornell *et al* 1995). Hydrogens were added to the model before minimisation could be performed. Structure minimization was performed in vacuum with all options set at the default value (steepest descent steps: 100, Steepest descent step size: 0.02, conjugate gradient steps: 100, conjugate gradient step size: 0.02, update interval: 10). Particular atoms could be fixed if required so that they would not be adjusted in the minimisation, for example if only atoms in a particular region were intended to be minimised, the rest of the atoms in the model could be fixed. Charges were assigned automatically to standard residues with AMBER ff99SB force field (default selection in Chimera minimise). Non-standard residues were assigned with AM1-BCC charge method (Jakalian *et al* 2000; 2002), and standard new charges were assigned to any ions present in the model (e.g. -1 for Cl^-, and +1 for Na^+). Charges were computed using ANTECHAMBER (Wang *et al* 2006).

2.5. Structural analysis of amino acid substitutions using homology models

Known non-synonymous missense mutations that result in an amino acid substitution were analysed within homology models in order to analyse the structural consequences of the substitution.

2.5.1. Replacing amino acids within a structure

Amino acids were substituted using the 'swapaa' command in Chimera, then the residue and the surrounding area within the model were thoroughly explored by examining all of the possible rotamers of the residue in the Dunbrack rotamer library (Dunbrack 2002), particularly focussing on clashes and hydrogen bonds with other surrounding residues and side-chains. Residues substituted as a result of a mutation can have various structural or functional effects within a protein that lead to a deleterious phenotype. The following list of structural analyses was compiled based on previous work (Martin *et al* 2002; Hurst *et al* 2009) and considered for each substitution.

- *Mutations affecting hydrogen bonding*: Hydrogen bonds are important for stabilizing the structure of a protein. If a residue involved in a hydrogen bond is replaced with a residue that is unable to form a hydrogen bond it is likely for the protein to become destabilized.
- *Mutations to proline*: Proline has a distinctive cyclic structure of its side chain that gives the amino acid exceptional rigidity. Substitution to a proline will often result in distortion of a structure. Proline residues will introduce kinks in α-helices and disrupt β-sheets.
- *Mutations from glycine*: Glycine is a very small amino acid with no side-chain, and is able to adopt conformations that would be sterically unachievable for other amino acids. Distortion will occur if a glycine in such a conformation is replaced with another amino acid.
- *Steric clashes*: Clashes of a substituted residue with surrounding residues will occur if it is larger than the native residue or has a larger side-chain, which will lead to distortion or incorrect folding.

- *Mutations within conserved regions*: Single residues or whole regions that are highly conserved between species are likely to have significant importance within the protein. Their conservation may be due to their importance in the structure or fold of the protein, interactions with ligands, ions or agonists or important interactions with other proteins.
- *Introducing a hydrophobic residue on the protein surface*: Hydrophobic residues are mostly found in the protein core, and can be unstable on the surface, due to charge. Substituting a surface residue to a hydrophobic amino acid could lead to protein aggregation.
- *Introducing a hydrophilic residue in the core*: Hydrophilic amino acids are generally found on the surface of a protein. The side chains of these amino acids are capable of hydrogen bonding, and therefore when introduced into the protein core, could destabilise the structure.
- *Disruption of buried charge*: Electrostatic interactions are very important in structural stability in some parts of proteins. If the net charge in such a region is disrupted by substitution of a positively charged residue with a negatively charged residue, or *vice versa*, the stability will become compromised and affect folding or function.
- *Cation-π interactions*: A cation-π interaction occurs between the face of an electron-rich aromatic side chain (phenylalanine, tyrosine or tryptophan) and a cationic side chain (lysine or arginine) (Gallivan *et al* 1999). They are comparable in strength to a hydrogen bond or salt bridge and can play an important role in stabilising the 3-dimensional structure.
- *π-π interactions*: π-π interactions (or π-stacking), occur between the faces of two electron-rich aromatic side chains leading to attractive, non-covalent interactions. Aromatic-aromatic interactions are described in the stacking of nucleotides in a strand of DNA (Hunter 1993). They are also important for the self-assembly of proteins and are prevalent in protein crystal structures (McGaughey *et al* 1998).
- *Disulphide bonds*: Disulphide bonds are also known as disulphide bridges, ss-bonds or cys-bonds. They are strong covalent bonds that occur by an oxidation reaction

between the thiol groups of two cysteine residues in close proximity and cannot exist in the reducing environment of the cell cytosol.

2.5.2. Finding non-bonded interactions, close contacts and clashes between atoms

The Chimera option 'FindHBond' identifies possible hydrogen bonds (H-bonds) using information about atom types, and geometric criteria based on a sample of small molecule X-ray crystal structures (Mills & Dean 1996). Possible H-bond donor groups are hydrogen-bearing nitrogen, and oxygen, and possible H-bond acceptor groups are nitrogen, and oxygen atoms with a lone pair. The surrounding area of each substituted amino acid was inspected for H-bonds with other atoms, for both the native and the mutant protein. Atoms within 6 Å of the substituted residue were selected. Chimera was instructed to calculate H-bonds in the selected surrounding region, and all potential H-bonding interactions fulfilling the criteria were shown.

The region surrounding the substituted residue was also inspected for possible cation-π interactions. The web server CaPTURE (http://capture.caltech.edu) uses criteria described by Gallivan and Dougherty in 1999 to identify energetically significant cation-π interactions within protein structures. The criteria, in brief, implement a variant of the OPLS (optimized potentials for liquid simulations) force field where only electrostatic and van der Waals interactions are considered, and only cation-π with binding energies above a particular threshold are reported (Gallivan & Dougherty 1999). If the structure was available on the PDB, then the PDB I.D. was submitted. If the structure was a homology model, then the coordinate file in pdb format was uploaded. The output was obtained within a few seconds reporting all possible cation-π interaction pairs and energetically significant pairs. These residues were checked to determine whether they fell within the region of the substituted amino acid of interest. Both the wild-type and the mutant protein coordinates were submitted to this server to determine whether any critical cation-π interactions were lost or gained as a result of the substitution.

The Chimera option 'Find Clashes/Contacts' is a useful tool that identifies interatomic clashes and contacts based on standard van der Waals (vdW) radii and user-specified criteria. Clashes are unfavourable interactions where atoms are too close in proximity to each other and in the case of a substituted amino acid, may lead to local disruptions in the protein. Contact refers to all kinds of direct interactions that can occur: polar and nonpolar, favourable and unfavourable (including clashes), salt-bridges, vdW forces (H-bonds are not included).

The Clash/Contact parameters dictate what is considered a clash or contact. The overlap between two atoms is defined as the sum of their vdW radii minus the distance between them and minus an allowance for potentially H-bonded pairs. Chimera identifies atoms with overlapping vdW radii and then reports them as either clashes or contacts depending on the overlap cut-offs specified (a larger positive cut-off restricts the results to more severe clashes, whereas a negative cut-off can identify favourable contacts). Default cut-off criteria were used when analysing both clashes and contacts (0.6 Å and -0.4 Å, respectively). For pairs of atoms that include a H-bond donor and acceptor, an allowance for a potential H-bond is subtracted. An allowance more than 0 Å reflects the observation that atoms sharing a hydrogen bond can come closer to each other than would be expected from their vdW radii and hence should not be identified as a clash. The default allowance criteria for clashes and contacts in such cases are 0.4 Å and 0 Å, respectively. The wild-type and the mutant proteins were checked using the same criteria to establish whether any clashes or contacts were introduced with the substituted amino acid, or whether any contacts were abolished.

Clashes and H-bonds were additionally assessed directly in the rotamer dialog. Once a list of rotamers had been calculated, clash and contact criteria was applied to each one in order to assess the predicted clashes and H-bonds with surrounding atoms for each rotamer of a single residue.

3. MOLECULAR ANALYSIS OF NOVEL GLYT2 MUTATIONS ASSOCIATED WITH STARTLE DISEASE

3.1. Background

Startle disease, also known as hyperekplexia, is a rare disorder characterized by evoked episodes of hypertonia. This neurological disorder can have serious consequences, including sudden infant death from apnoea episodes or aspiration pneumonia (Nigro & Lim 1992; Vigevano *et al* 1989). Although symptoms often diminish during the first year of life, the exaggerated startle response can persist into adulthood, leading to unprotected falls (Bakker *et al* 2006; Thomas *et al* 2010). Fortunately, hyperekplexia can be treated using the benzodiazepine clonazepam (Harvey *et al* 2008), which potentiates inhibitory $GABA_A$ receptor function. A physical intervention called the Vigevano manoeuvre, involving flexion of the head and limbs towards the trunk, also dissipates and counteracts the effects of acute hypertonia and apnoea episodes in newborns (Vigevano *et al* 1989). Hyperekplexia is now well recognized and often detected within specialist neurology and paediatric centres worldwide (Harvey *et al* 2008; Davies *et al* 2010). Collaborations between clinicians and research scientists have revealed that the primary cause of startle disease is defective inhibitory glycinergic transmission. Currently, the major genetic cause of hyperekplexia is missense, nonsense, frameshift or splice-site mutations in the postsynaptic glycine receptor (GlyR) α1 gene (*GLRA1*) (Shiang *et al* 1993, 1995; Chung *et al* 2010), although large deletions in this gene are also common in patients of Kurdish descent (Becker *et al* 2006; Sirén *et al* 2006). In a handful of cases, mutations in the genes encoding the GlyR β subunit (*GLRB*) (Rees *et al* 2002; Al-Owain *et al* 2012; Lee *et al* 2012) and the synaptic clustering proteins gephyrin (*GPHN*) (Rees *et al* 2003) and collybistin (*ARHGEF9*) (Harvey *et al* 2004) have also been linked to hyperekplexia. However, mutations in the latter two genes are more commonly associated with MOCO deficiency (Reiss & Hahnewald 2011) and X-linked intellectual disability (Kalscheuer *et al* 2010; Holman *et al* 2012).

Since many remaining cases do not harbour defects in these genes, some years ago researchers began to explore presynaptic causes of startle disease. This led to the

identification of missense, nonsense and frameshift mutations in the GlyT2 gene (*SLC6A5*), encoding a Na^+/Cl^--dependent neurotransmitter transporter that maintains a high presynaptic pool of glycine at glycinergic synapses (Eulenburg *et al* 2006; Rees *et al* 2006). Detailed structure-function analyses revealed that GlyT2 mutations disrupted transporter membrane trafficking, Na^+ or glycine binding sites (Rees *et al* 2006). Subsequently, unique GlyT2 mutations were discovered in cattle and dogs, causing startle disorders with early neonatal lethality (Charlier *et al* 2008; Gill *et al* 2011).

3.2. Study aims

I participated in an international screening program for GlyT2 mutations in startle disease, in collaboration with Prof. Mark Rees at the Institute of Life Science, University of Swansea. Resequencing of *SLC6A5* in 93 unrelated hyperekplexia probands identified 20 new recessive mutations within 17 index cases, of which 19 are novel variants, and nine were missense mutations (Carta *et al* 2012). An additional collaborative study was carried out with Beatriz López-Corcuera at the Departamento de Biología Molecular Centro de Biología Molecular 'Severo Ochoa', Madrid, Spain. Here, a novel dominant missense mutation (Y705C) was identified in eight individuals from Spain and the UK (Giménez *et al* 2012). My specific aims within these studies were:

- To use site-directed mutagenesis to introduce missense mutations into a human GlyT2 expression construct in the vector pRcCMV for use in [^{3}H]glycine uptake assays.
- To build a homology model of human GlyT2 based on the crystal structure of the *Aquifex aeolicus* leucine transporter.
- To use this homology model to predict how GlyT2 mutations specifically disrupt transporter structure and function.
- To correlate the results of [^{3}H]glycine uptake assays with molecular modelling predictions to determine the likely pathogenic mechanisms of GlyT2 mutations.

Results

3.3. Identification of new mutations in the human GlyT2 gene (*SLC6A5*) in startle disease

DNA samples from individuals with a clinical diagnosis of startle disease were ascertained by referral from neurologists, paediatricians or clinical geneticists from international centres. Informed consent and clinical data were obtained by the referring clinician (Local Research Ethics Committee, the South West Wales REC). Clinical criteria included a non-habituating startle response (positive nose-tap test), history of neonatal/infantile hypertonicity and negative results in *GLRA1* gene screening by Sanger sequencing and MLPA (Multiplex Ligation-Dependent Probe Amplification). Exons and intron-exon boundaries of *SLC6A5* were amplified by PCR from genomic DNA isolated from peripheral blood of patients using established primers sets (Rees *et al* 2006). PCR products were directly sequenced following gel electrophoresis and purification.

Sequence variants were assigned as potentially disease-causing mutations after exclusion from a panel of 400 human controls and cross-referencing with common *SLC6A5* polymorphisms (Rees *et al* 2006) and those found in dbSNP (http://www.ncbi.nlm.nih.gov/projects/SNP/). This revealed 20 sequence variants that were found exclusively in 17 index cases, 14 of which showed homozygous or compound heterozygous recessive inheritance (Fig. 3.1; Table 3.1; Carta *et al* 2012). These included four nonsense mutations (W151X, R191X, Y297X, R439X), four frameshift mutations (P108L+fs25X, L198R+fs123X, S489F+fs39X, I665K+fs1X) and three splice site mutations (IVS14+1ΔG, IVS13+1 G>T, IVS8+1 G>A). Nine novel missense mutations affected residues in TM2 (L237P, P243T and E248K), TM3 (A275T), TM7 (S513I), the extracellular TM7-TM8 loop (F547S) and the intracellular TM10-TM11 loop (Y656H, G657A). These missense mutations were assessed using the bioinformatics programmes SIFT (http://sift.jcvi.org/), which uses sequence conservation (Ng 2003), and Polyphen-2 (http://genetics.bwh.harvard.edu/pph2), which uses both sequence conservation and structural predictions (Adzhubei *et al*

2010), to assess whether an amino acid substitution is likely to affect protein function. The predictions from these two servers are summarised in Table 3.2.

Case	Inheritance	Exon	Genotype	Consequences	Class
1	AR	14	IVS14+1ΔG (H)	Impact on splicing	Splice site
2*	AR	4	C727A (H)	P243T	Missense
3*	AR	5	C891A (H)	Y297X	Nonsense
4	AR	8	C1315T (H)	R439X	Nonsense
5	AR	8	C1315T (H)	R439X	Nonsense
6	AR	8	C1315T (H)	R439X	Nonsense
7	CH	8 13	C1315T (P) IVS13+1 G>T (M)	R439X Impact on splicing	Nonsense Splice site
8†	CH	11 13	T1640C (M) T1966C (†)	F547S Y656H	Missense Missense
9*	CH	4 8	G742A (P) IVS8+1 G>A (M)	E248K Impact on splicing	Missense Splice site
10‡	NC	9 14	ΔCT [1460-1467] (‡) G1970C (‡)	S489F+fs39X G657A	Frameshift Δ Missense
11*	CH	3 3	C571T (P) ΔTG [593-594] (M)	R191X L198R+fs123X	Nonsense Frameshift Δ
12	NC	2	G452A (P)	W151X	Nonsense
13	NC	8	C1315T (P)	R439X	Nonsense
14	NC	5	G823A (P)	A275T	Missense
15	CH	2 14	ΔC [319-323] (P) ΔT [1994] (M)	P108L+fs25X I665K+fs1X	Frameshift Δ Frameshift Δ
16	CH	3 4	ΔTG [593-594] (P) T710C (M)	L198R+fs123X L237P	Frameshift Δ Missense
17	AR	10	G1538T (H)	S513I	Missense

Table 3.1. Details of *SLC6A5* hyperekplexia mutations.

Key: *Indicates an affected sibling with the same mutation(s), †indicates where only one parent was available to assess compound heterozygote status since the maternal allele did not carry the Y656H mutation. ‡Indicates parental samples were lost to contact and long-range PCR attempts were unsuccessful (22 kb target). Mode of Inheritance: NC - not confirmed; CH - compound heterozygosity; AR - autosomal recessive. Genotype: Homozygous recessive inheritance is indicated by (H) and for compound heterozygosity: (P) = paternal allele, (M) = maternal allele. Δ = deletion.

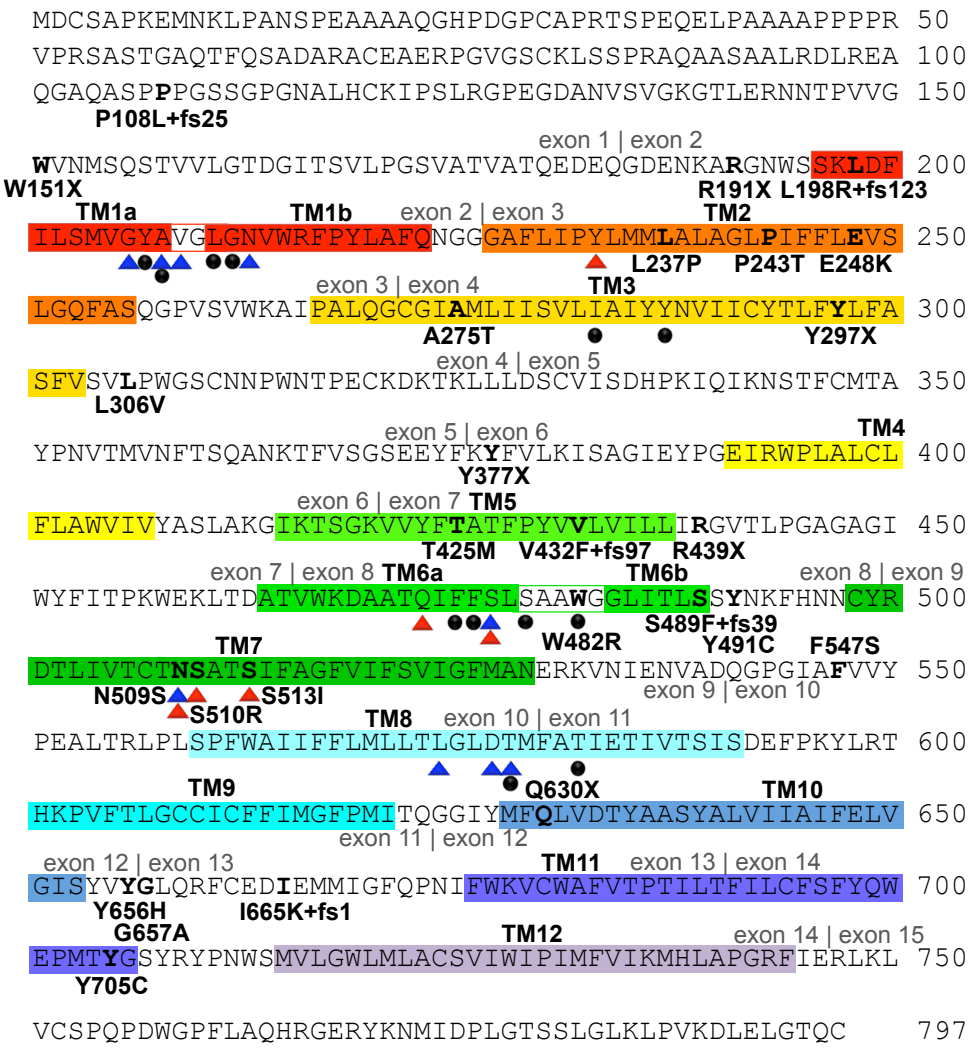

Figure 3.1. New human GlyT2 mutations identified in individuals with startle disease. Amino acid sequence of human GlyT2 (UniProt entry: Q9Y345) indicating the positions of putative transmembrane (TM) domains (coloured boxes) and amino acid residues affected by hyperekplexia mutations (Eulenburg *et al* 2006; Rees *et al* 2006; Carta *et al* 2012). Blue and red triangles indicate residues in GlyT2 that are likely to co-ordinate Na^+ and Cl^- ions, respectively, based on structure/function studies of the bacterial leucine transporter LeuT and other mammalian neurotransmitter transporters such as GAT-1 and SERT. However, GlyT2 binds three Na^+ ions (Supplisson & Roux 2002), while LeuT binds two, suggesting that other residues involved in Na^+ co-ordination still remain to be identified. Filled black circles indicate residues predicted to be involved in glycine binding.

An additional dominant mutation (c.2114A>G, Y705C) was found in exon 15 of *SLC6A5* in eight individuals from three families in two cohorts of hyperekplexia patients that were devoid of GlyR gene mutations (Giménez *et al* 2012).

Position	Exon	Protein	Type	SIFT	PolyPhen-2	Overall Prediction
c.T710C	4	L237P	Missense	Damaging (0)	Probably damaging (1.0)	Pathogenic (Carta *et al* 2012)
c.C727A	4	P243T	Missense	Damaging (0)	Probably damaging (1.0)	Pathogenic (Carta *et al* 2012)
c.G742A	4	E248K	Missense	Damaging (0)	Probably damaging (1.0)	Pathogenic (Carta *et al* 2012)
c.G823A	5	A275T	Missense	Tolerated (0.3)	Possibly damaging (0.774)	Pathogenic (Carta *et al* 2012)
c.C916G	5	L306V	Missense	Damaging (0.02)	Probably damaging (1.0)	Pathogenic (Rees *et al* 2006)
c.C1274T	8	T425M	Missense	Damaging (0)	Probably damaging (1.0)	Pathogenic (Rees *et al* 2006)
c.T1444C	9	W482R	Missense	Damaging (0)	Probably damaging (1.0)	Pathogenic (Rees *et al* 2006)
c.A1526G	10	N509S	Missense	Damaging (0)	Possibly damaging (0.998)	Pathogenic (Rees *et al* 2006)
c.T1530G	10	S510R	Missense	Damaging (0.01)	Possibly damaging (0.935)	Pathogenic (Rees *et al* 2006)
c.G1538T	10	S513I	Missense	Damaging (0)	Probably damaging (1.0)	Pathogenic (Carta *et al* 2012)
c.T1640C	11	F547S	Missense	Damaging (0)	Probably damaging (1.0)	Pathogenic (Carta *et al* 2012)
c.T1966C	13	Y656H	Missense	Damaging (0)	Probably damaging (1.0)	Pathogenic (Carta *et al* 2012)
c.G1970C	14	G657A	Missense	Damaging (0)	Possibly damaging (0.931)	Pathogenic (Carta *et al* 2012)
c.A2114G	15	Y705C	Missense	Damaging (0)	Probably damaging (1.0)	Pathogenic (Giménez *et al* 2012)

Table 3.2. Single nucleotide variants in *SLC6A5* causing hyperekplexia. Prediction of pathogenicity of the likelihood of missense mutations leading to amino acid substitutions found in individuals with hyperekplexia using the SIFT (http://sift.jcvi.org/www/SIFT_enst_submit.html) and PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/index.shtml) web servers. All mutations are predicted to be damaging with over 90% confidence by both servers, with the exception of A275T of which SIFT predicts to be tolerated and PolyPhen-2 predicts to be possibly damaging but with only 77% confidence.

Position	Exon	Protein Precursor/mature	Type	SIFT	PolyPhen-2	Overall Prediction
c.T596G	6	p.M199R / p.M177R	Missense	Damaging (0)	Possibly damaging (0.802)	Pathogenic (Al-Owain *et al* 2012)
c.G752A IVS5+5G>A	8 -	p.G251D / p.G229D -	Missense Splice site	Damaging (0)	Probably damaging (1.0) -	Pathogenic (Rees *et al* 2002) Pathogenic (James *et al* 2012)
c.920_921D InsGA	9	p.L307R / p.L285R	Missense	Damaging (0)	Probably damaging (0.997)	Pathogenic (James *et al* 2012)
c.G996T	9	p.W332C / p.W310C	Missense	Damaging (0)	Probably damaging (0.999)	Pathogenic (James *et al* 2012)
c.C1028T	9	p.S343F / p.S321F	Missense	Damaging (0)	Probably damaging (0.990)	Pathogenic (Lee *et al* 2012)

Table 5.1. Single nucleotide variants in *GLRB* causing hyperekplexia. Prediction of pathogenicity of missense mutations in *GLRB* leading to amino acid substitutions found in individuals with hyperekplexia using the SIFT (http://sift.jcvi.org/www/SIFT_enst_submit.html) and PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/index.shtml) web servers.

3.4. [^{3}H]glycine uptake assays demonstrate that the majority of GlyT2 mutations impair glycine uptake

Mutations were introduced into an expression construct encoding human GlyT2 (pRcCMV-hGlyT2; Rees *et al* 2006) using the QuikChange site-directed mutagenesis kit (Stratagene) and the complete coding region was sequenced to verify that only the desired mutation had been introduced. To examine the effects of GlyT2 mutations on glycine uptake, HEK293 cells were transfected with wild-type and mutant GlyT2 constructs and used for [^{3}H]glycine uptake assays (Rees *et al* 2006; Carta *et al* 2012). GlyT2 constructs harbouring nonsense (W151X, R191X, Y297X, R439X), frameshift (P108L+fs25, L198R+fs123, S489F+fs39, I665K+fs1) and missense (L237P, P243T, S513I, F547S, Y656H, G657A) mutations revealed no significant [^{3}H]glycine uptake in HEK293 cells compared to control values (Fig. 3.2). By contrast, the A275T mutant was capable of [^{3}H]glycine uptake, although at a lower level than wild-type GlyT2. Further characterization of the A275T mutation in *Xenopus laevis* oocytes by Stéphane Supplisson (Ecole Normale Supérieure, Paris, France) demonstrated that this mutation results in a decrease in the affinity of Na^+ binding and an increase in glycine EC_{50} at depolarizing potentials, which reduces the ability of the transporter to pump glycine at low concentrations, a situation that prevails at the synaptic cleft (Carta *et al* 2012).

[^{3}H]glycine uptake assays for the dominant Y705C mutation were carried out by the López-Corcuera group using COS7 cells. The V_{max} of glycine transport by the Y705C mutant was around 60% of the wild-type GlyT2 V_{max} (23 ± 2.5 for GlyT2^{Y705C} versus 38 ± 3.2 nmol gly/mg protein/5 min for wild-type). The co-expression of wild-type and Y705C transporters resulted in a V_{max} of 30 ± 3.1 nmol gly/mg protein/5 min, which results in around 79% of wild-type GlyT2.

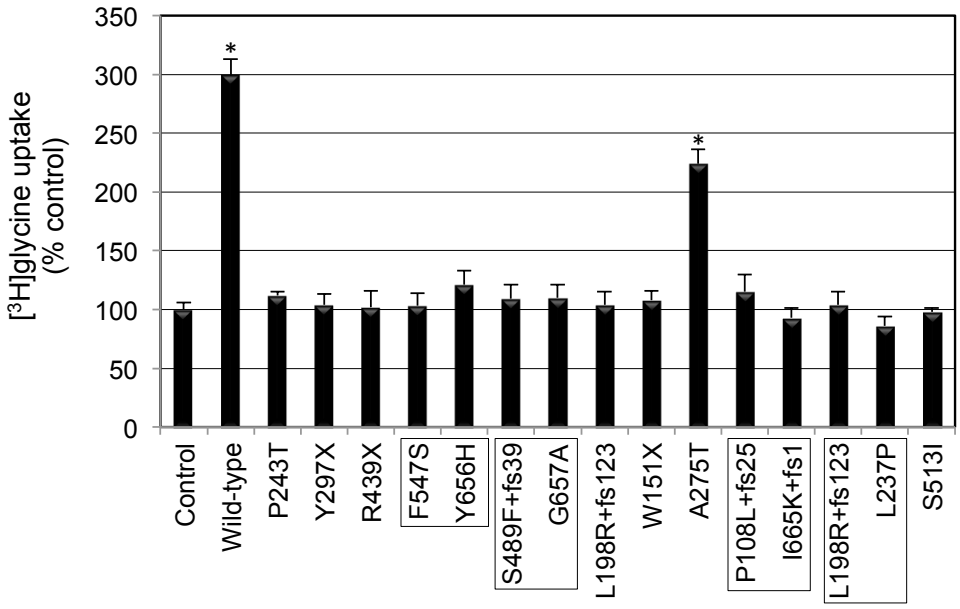

Figure 3.2. Functional activity of human GlyT2 and hyperekplexia mutants. Glycine uptake in HEK293 cells transiently expressing hGlyT2 and single hyperekplexia mutants after 5 min incubation with [^{3}H]glycine (60 Ci/mmol, NEN) at a final concentration of 300 µM (Carta *et al* 2012). Since low levels of glycine uptake are found in HEK293 cells (Rees *et al* 2006), [^{3}H]glycine uptake was calculated as nmoles/min/mg protein then expressed as a percentage of the empty expression vector transfected control. Only wild-type GlyT2 and mutant A275T results in any detectable uptake above control values.

3.5. Building a homology model of human GlyT2

Fold recognition of human GlyT2 was performed with GenTHREADER (Lobley *et al* 2009) and HHPred (Söding *et al* 2005). The structure of the bacterial leucine transporter (LeuT; PDB: 2A65; Yamashita *et al* 2005) was identified as the best template using both approaches (GenTHREADER p-value: 2e-21; HHPred e-value: 0). Profile-profile alignments were generated between the human GlyT2 sequence and the LeuT structure using MUSCLE (http://www.ebi.ac.uk/Tools/muscle/index.html; Edgar 2004) and T-coffee ('Combine' option) web servers (http://www.ebi.ac.uk/Tools/t-coffee/) (Notredame *et al* 2000). The final alignment between GlyT2 and LeuT was visually checked in the molecular graphics program Chimera (Petterson *et al* 2004) comparing the actual secondary structure of the template structure with the predicted secondary structure of GlyT2. Secondary

structure for the GlyT2 sequence was predicted using TMHMM server, a server designed to predict transmembrane helices in proteins using a hidden Markov model (http://www.cbs.dtu.dk/services/TMHMM/; Krogh *et al* 2001). The alignment was visually assessed in Chimera and changes were made manually by moving residues that were not optimally aligned based on the actual and predicted secondary structures. The final alignment was prepared in the 'pir' format so that MODELLER could read it and is shown in Fig. 3.3.

A large part of the extracellular loop between TM3 and TM4 (residues 312-354) was removed since it has little or no sequence identity with the template or any other known structure and therefore was not possible to model this region with any accuracy. Full stops indicate heteroatoms, which in this case are the two Na^+ ions. GlyT2 is known to bind three Na^+ ions (Roux & Supplisson 2000), but the binding site of one of these Na^+ ions remains elusive since there are currently no structural homologues available that bind more than two Na^+ ions. The 'L' and the 'G' at the end of the sequences indicate the bound ligands leucine and glycine, respectively so that MODELLER would include a bound glycine and two Na^+ ions within the model based on the coordinates of the leucine and Na^+ ions in the LeuT template structure.

The final alignment resulted in 26% sequence identity between GlyT2 and LeuT, and 38% sequence similarity calculated using the sequence identities and similarities (SIAS) server (http://imed.med.ucm.es/Tools/sias.html) where amino acids were grouped by aromaticity, charge, polarity and size. Based on this alignment, 50 homology models of human GlyT2 were generated using MODELLER-9v7 (Eswar *et al* 2008) and assessed with the DOPE statistical potential score (Shen & Sali 2006). The model with the lowest score was selected, and the normalized DOPE Z-score was calculated for the model.

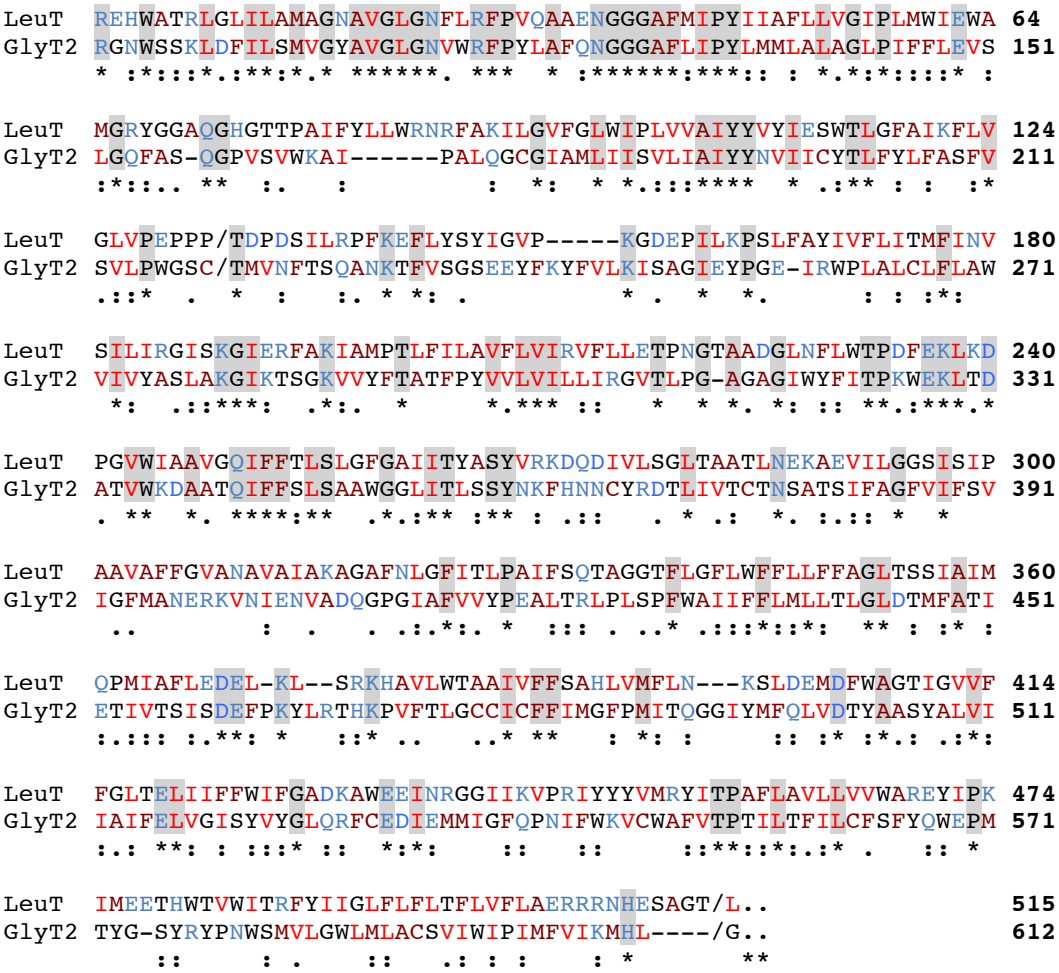

Figure 3.3. Alignment of LeuT and GlyT2 sequences. Residues are coloured according to their hydrophobicity using the Kyte-Doolittle hydrophobicity scale (Kyte & Doolitte 1982), where the most hydrophobic residues are coloured in bright red (Iso, Val, Leu), then less hydrophobic in dark red (Phe, Cys, Ala, Met) and black (Gly, Thr, Trp, Ser, Tyr, Pro). Hydrophilic and charged residues are coloured in blue (Asp, Glu, Asn, Gln, Arg, His, Lys). Residues highlighted in grey are conserved between LeuT and GlyT2 (also denoted by a *) and residues that are partially conserved (grouped by similar residue characteristics, such as charge, aromaticity, polarity and size) are indicated with a colon or stop. Gaps are indicated with a hyphen (-) and a forward slash (/) indicates a chain break.

A normalized DOPE Z-score is used to assess the overall quality of a homology model by normalising the raw DOPE scores of the model over its sequence length. This gives some assessment of model, where positive scores are likely to represent poor models and negative scores, particularly below -1 are likely to represent a native-like structure. The GlyT2 homology model was calculated to have a normalized DOPE Z-score of -

1.518, indicating a reasonable model, particularly considering the use of a template structure with less than 30% sequence identity. The DOPE score of each residue of the selected model was plotted alongside the corresponding residues in the LeuT template structure (Fig. 3.4). Peaks of high energy, particularly DOPE scores that are more than 0, are considered unfavourable. Areas in the plot where the energies overlap are usually correlated with regions of high conservation between target and template. This step was useful for highlighting specific regions of a model with high energies and therefore likely inaccuracies. Generally, regions of particularly high energies were in agreement with those of LeuT, suggesting a native-like structure and a good model (Fig. 3.4).

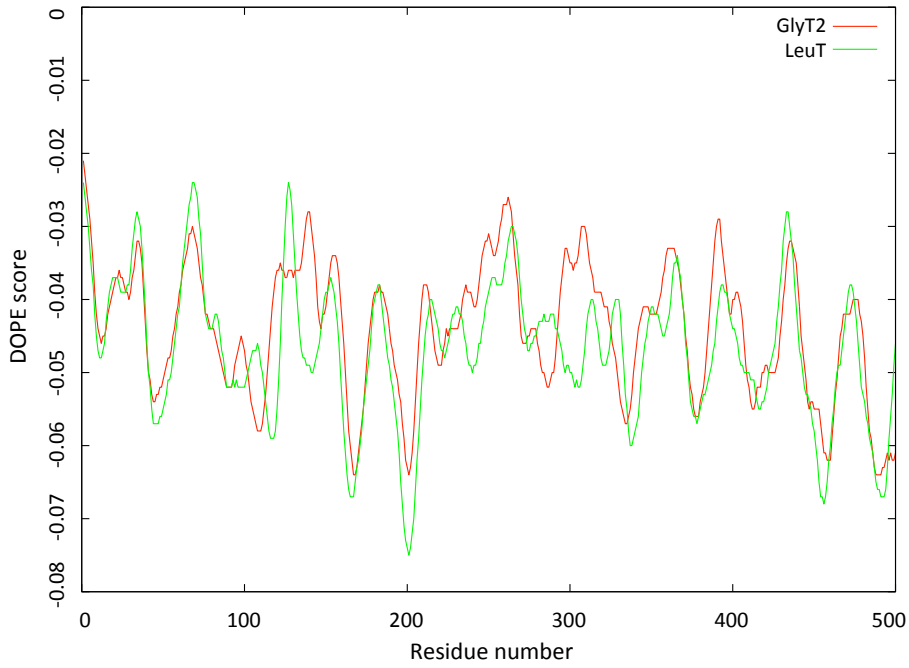

Figure 3.4. DOPE energy profiles of the human GlyT2 model and the LeuT structure. The DOPE score for each residue of both the GlyT2 homology model and the LeuT crystal structure plotted against the sequential residue number (1-500).

Additional evaluation of the homology model was performed using the QMEAN web server (http://swissmodel.expasy.org/qmean/cgi/index.cgi) (Benkert *et al* 2008, 2009). The QMEAN6 score of the model (Z-score = -4.18) fell within a range typically found for native proteins of similar size (Fig. 3.5) (Benkert *et al* 2009, 2011). Membrane proteins

are actually expected to have low QMEAN Z-scores, since the scoring function is based on a sample of proteins taken from PDB, of which most are globular proteins. Membrane proteins are significantly more difficult to crystal and therefore there are a much smaller number of membrane proteins in the PDB relative to soluble ones. There are around 83,000 proteins in the PDB, of which only around 600 of them are membrane proteins according to SCOP, the Structural Classification of Proteins database (http://scop.mrc-lmb.cam.ac.uk/scop/).

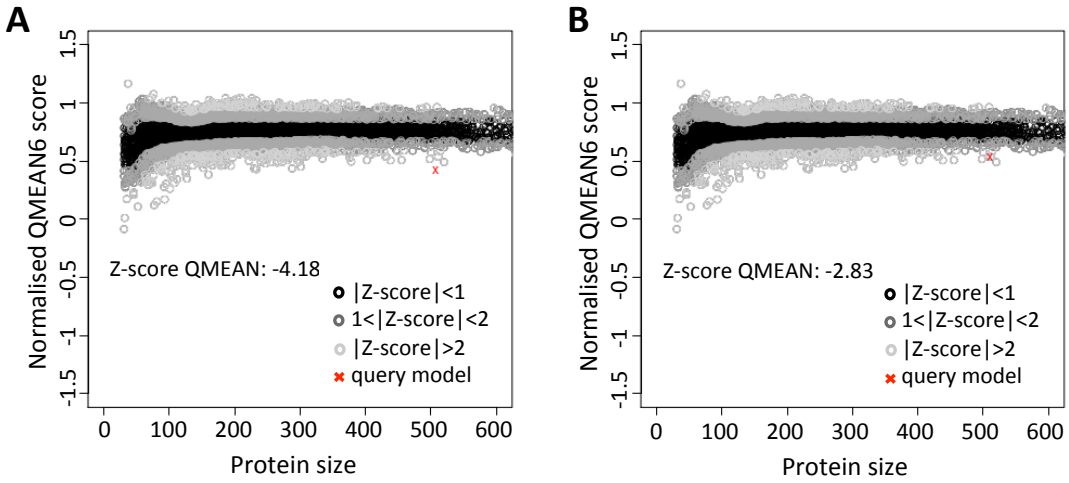

Figure 3.5. QMEAN scores of GlyT2 and LeuT model plotted against scores for other native proteins. Normalised QMEAN6 scores for human GlyT2 homology model (**a**) and LeuT structure (**b**) plotted alongside a sample of other scores for crystal structures of other native proteins. The score for the GlyT2 model lies just outside of the cluster of scores for proteins of a similar size, where the score for the LeuT structure lies within the cluster but on the border. This indicates that the GlyT2 homology model has a reasonably good score since it is close to the score of the native structure template, LeuT.

3.6. Building a molecular model of the GlyT2 Cl^- binding site

For modelling the GlyT2 chloride-ion binding site, Cl^- was placed into our model at the coordinates of the carboxylate carbon of the LeuT E290 side chain (Zomot *et al* 2007), based on the assumption that this position may be occupied by Cl^- in homologous transporters (Forrest *et al* 2007; Zomot *et al* 2007; Ben-Yona *et al* 2011). The residues thought to be involved in the coordination of Cl^- ion are generally conserved within the

SLC6 transporter family. In my model, these are: Y233 in TM2, Q473 and S477 in TM6, N509 and S513 in TM7 (indicated by red triangles in Fig 3.1), which correspond to Y47, Q250, T254, N286 and E290 in LeuT, respectively. Energy minimization was carried out locally in vacuum on the model with the *minimize structure* command in Chimera, using the Molecular Modelling ToolKit (MMTK) method (Hinsen 2000) and the AMBER94 force field (Cornell *et al* 1995) allowing Cl^- to fit more favourably in the binding pocket. In the resulting minimized structure, all distances between Cl^- and the side chains of the interacting residues are below 4.5 Å and therefore in a close enough range to interact.

3.7. Molecular modelling analysis of startle disease missense mutations in GlyT2

Alignment of the *Aquifex aeolicus* LeuT with GlyT2 allowed residues potentially involved in coordinating glycine or Na^+ and Cl^- binding to be identified. Substitution of Na^+ and glycine binding residues in GlyT2 has previously been shown to underlie pathogenicity in startle disease (Rees *et al* 2006). However, many of the novel GlyT2 mutations discovered in this study did not lead to substitutions of residues that directly bind Na^+ or glycine. To gain further insight into the precise molecular mechanisms underlying these new missense mutations, I examined each mutation individually in the context of the homology model of GlyT2. Amino acids were substituted using the '*swapaa*' command in Chimera (Pettersen *et al* 2004), then the residue and the surrounding area within the model was thoroughly explored by examining all possible rotamers of the residue side chain in the Dunbrack rotamer library (Dunbrack 2002), taking into account the lowest clash score, highest number of H-bonds and highest rotamer probability. A pragmatic approach was taken into examining each of the mutations by carrying out each of the structural analyses described in the methods of this thesis (see also Martin *et al* 2002; Hurst *et al* 2009) for each mutation.

Three of the missense mutations result in substitution of residues in TM2 (L237P, P243T and E248K). Introduction (L237P) or elimination (P243T) of proline residues in

transmembrane helices is predicted to destabilize the native fold of the helix, thus resulting in accelerated degradation or aggregation of the mutant protein (Bajaj *et al* 2007). By contrast, substitution E248K in TM2 results in the replacement of a negatively charged glutamate residue for a positively charged lysine, with a longer side chain that introduces potential clashes with A480 and A481 in TM6 (Fig. 3.6b,c). These clashes, particularly those with the backbone atoms of both A480 and A481, may disrupt the local fold of this unwound region of TM6 surrounding the glycine-binding site and affect the orientation of the key glycine-binding residue W482.

Substitution A275T in TM3 results in predicted clashes with I279, which may indirectly alter the orientation of the predicted glycine-binding residue Y287 further along TM3 (Fig. 3.6d,e). The corresponding residue in LeuT (Y108) anchors the leucine substrate, and also stabilizes the unwound region of TM1 (Singh *et al* 2008) by binding the amide nitrogen of L25 - equivalent to L211 in GlyT2. Thus, A275T could indirectly destabilize the unwound region of TM1, which contains several determinants of glycine and Na^+ binding.

Substitution S513I in TM7 introduces a much larger side chain at this position, resulting in predicted clashes with N213 (TM1) and N509 (TM7), both of which are predicted to be involved in Na^+ binding (Fig. 3.6f,g). It is also noteworthy that the recent discovery of residues involved in Cl^- binding in GAT-1 and SERT suggests that residues Y233 (TM2), Q473 and S477 (TM6) together with N509 and S513 (TM7) are likely to perform equivalent roles in GlyT2 (Fig. 3.7a) (Forrest *et al* 2007; Zomot *et al* 2007; Ben-Yona *et al* 2011). The substitution of serine for isoleucine at this position introduces a larger side chain (Fig. 3.7a,b), which lacks the hydroxyl necessary for the coordination of the Cl^- ion and may prevent Cl^- binding. Since a hydroxyl at this relative position is highly conserved in SLC6 transporters (Forrest *et al* 2007; Zomot *et al* 2007; Ben-Yona *et al* 2011), the S513I substitution is predicted to interfere with glycine transport by disrupting interactions with both Cl^- (directly) and Na^+ (indirectly).

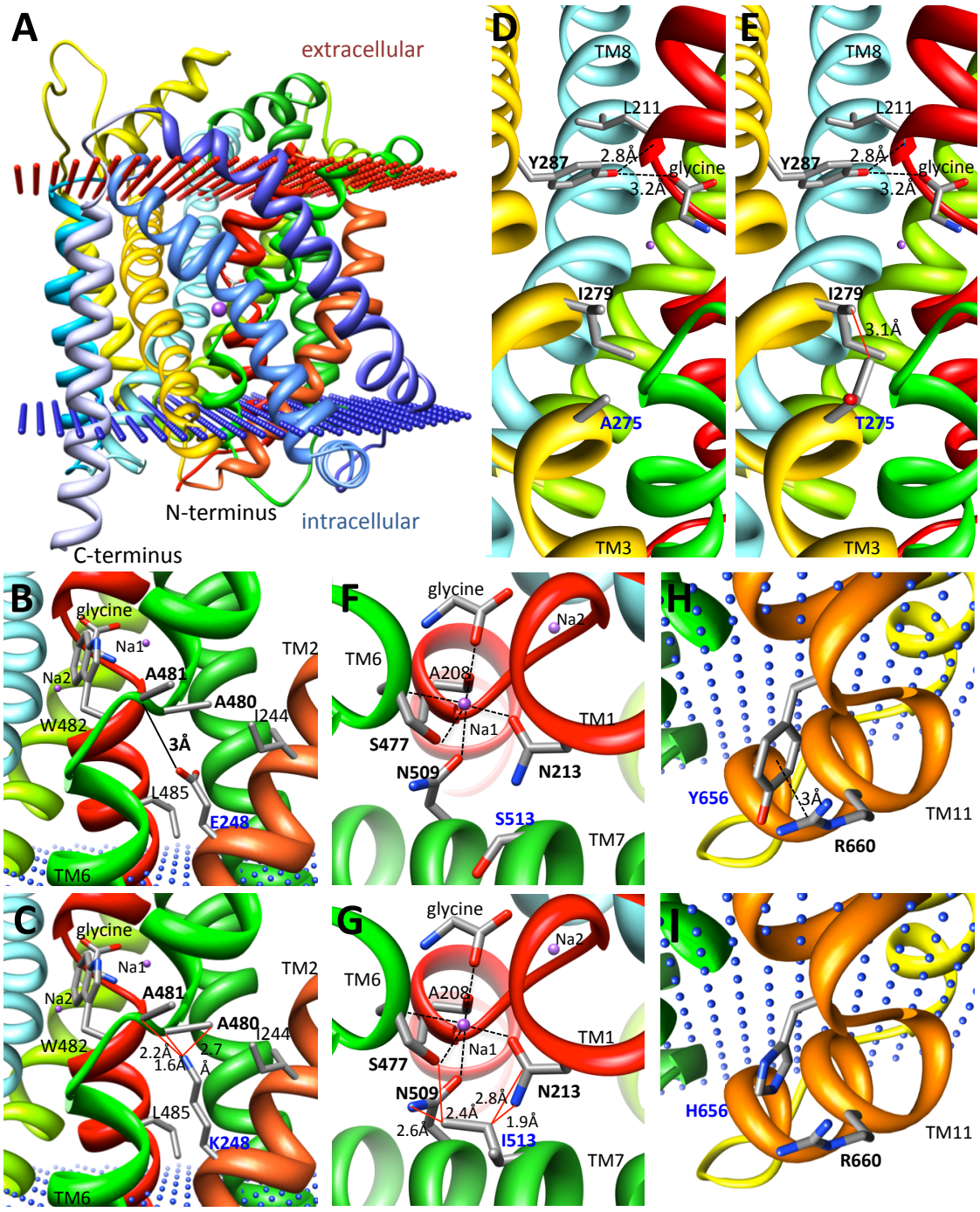

Figure 3.6. Molecular modelling of GlyT2 mutations. A: Side view of the human GlyT2 monomer showing coloured transmembrane helices as ribbons. Glycine and two (of three) Na^+ ions (purple spheres) are depicted. Note that the TM3-TM4 extracellular loop (EL2; indicated on Fig 3.8a) was not modelled due to an insertion of residues 312-354 in GlyT2 relative to LeuT. **B** and **C**: Substitution E248K in TM2 replaces a negatively-charged glutamate for a positively-charged lysine, with a longer side chain that introduces potential clashes (red lines) with A480 and A481 in TM6. This is likely to affect the orientation of the key glycine-binding residue W482. **D** and **E**: Substitution

A275T in TM3 results in several predicted clashes with I279 that are predicted to alter the orientation of the potential glycine-binding residue Y287 and destabilize the unwound region of TM1, which contains several determinants of glycine and Na^+ binding. **F** and **G**: Substitution S513I in TM7 introduces a larger side chain at this position, resulting in predicted clashes with N213 (TM1) and N509 (TM7), both of which are involved in the binding of Na^+. Note that N509 and S513 are also predicted to be involved in coordinating Cl^- binding to GlyT2 (see figure 3.7). **H** and **I**: Substitution Y656H in the intracellular TM10-TM11 loop is predicted to abolish a cation-π interaction between Y656 and R660 that may affect folding, stability or interfere with possible intracellular GlyT2 accessory protein interactions.

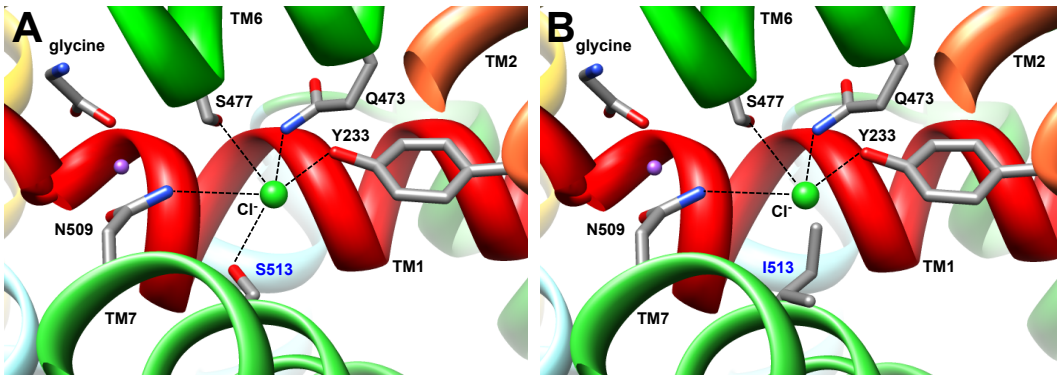

Figure 3.7. Molecular modelling of the GlyT2 chloride-ion binding site. A: Magnified view of the human GlyT2 chloride ion-binding site showing Cl^- as a green sphere and coordination by Y233 in TM2, Q473 and S477 in TM6, N509 and S513 in TM7. **B**: Substitution S513I in TM7 introduces a larger side chain, which lacks the essential hydroxyl necessary for the coordination of Cl^- and may occlude the binding site. Since a hydroxyl at this position is highly conserved in SLC6 transporters (Forrest *et al* 2007; Zomot *et al* 2007; Ben-Yona *et al* 2011), this substitution is likely to disrupt glycine uptake by directly interfering with coordination of Cl^-.

Substitution F547S affects a highly conserved residue found in extracellular TM7-TM8 loop 4B (EL4B). EL4B is predicted to be composed of two short helices separated by an acute bend and participates in conformational mobility during the transport cycle in DAT (Norregaard *et al* 1998), GAT-1 (Zomot *et al* 2003), GlyT1 (Ju *et al* 2004) and SERT (Mitchell *et al* 2004). Substitution Y656H in the intracellular TM10-TM11 loop is predicted to abolish a cation-π interaction between Y656 and R660 (Fig. 3.6h,i). This interaction is likely to have a significant role in the folding or stabilization of the

protein in this unwound region. The substitution may also interfere with important intracellular GlyT2 accessory protein interactions. The neighbouring residue was substituted (G657A) in another individual with hyperekplexia, which is predicted to have a similar mechanism by somehow abolishment of a cation-π interaction or interference with interactions with other proteins.

Lastly, the substitution Y705C results in the loss of a predicted hydrogen bond with the side chain of E701 (Fig 3.9). Y705 is a highly conserved residue between GlyT2 in different species (Fig. 3.8, upper panel) as well as within the superfamily of SLC6 transporters (Fig. 3.8, lower panel).

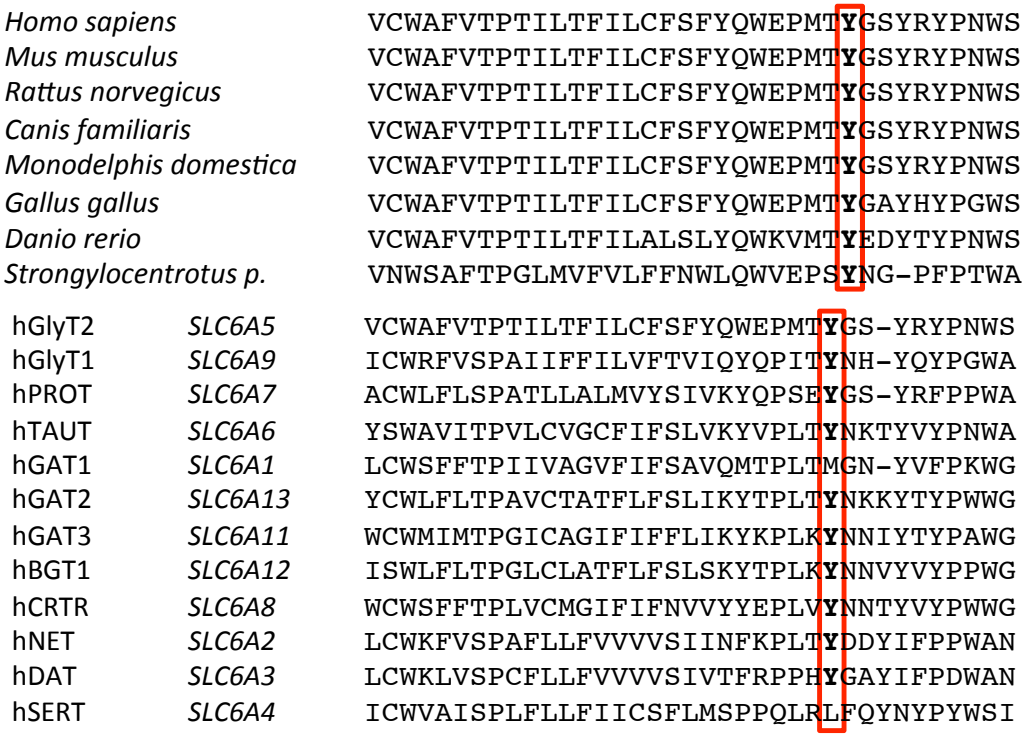

Figure 3.8. Phylogenetic and familial alignments of TM11 in GlyT2. Upper panel: Phylogenetic comparison of TM11 of GlyT2 containing Y705 (boxed). Lower panel: Sequence alignment of GlyT2 TM11 with other proteins from the human *SLC6* superfamily. Sequences were obtained from NCBI (www.ncbi.nlm.nih.gov) and were aligned using EBI MUSCLE alignment server. Tyrosine is found at the equivalent position in most human neurotransmitter transporters from the GABA, amino acid and monoamine subfamilies (NET, DAT) with the exception of GAT1 and SERT.

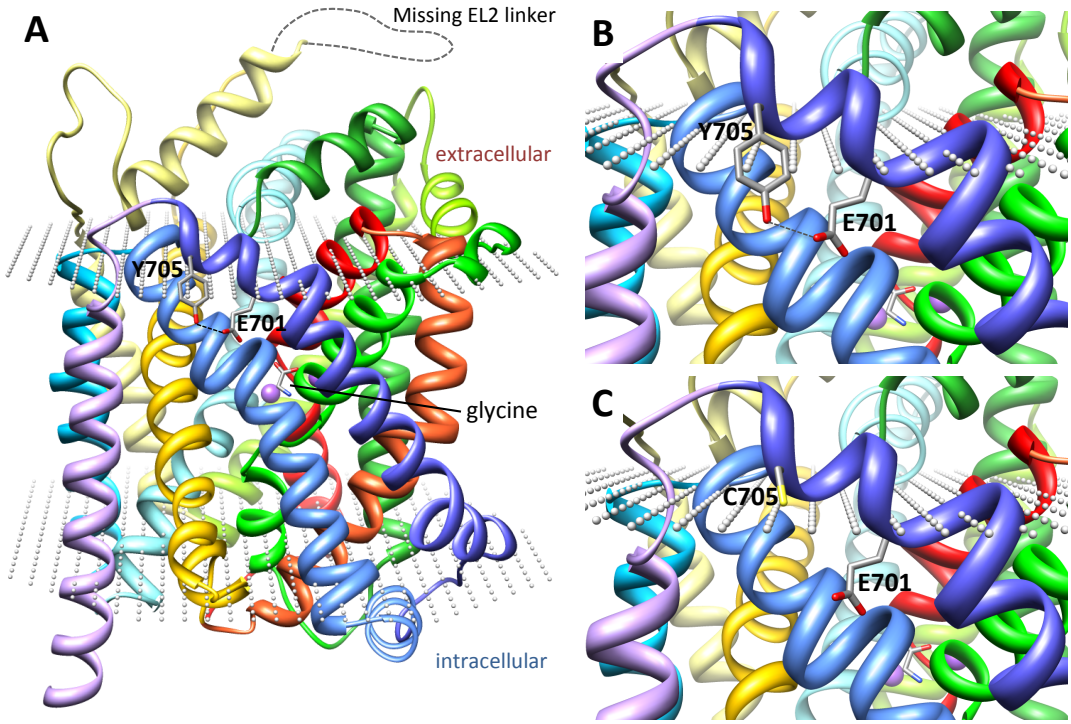

Figure 3.9. Molecular model of GlyT2 showing the localization of Y705 in TM11. A: Homology model of the GlyT2 monomer showing the predicted position of Y705 at the membrane interface. **B**: Y705 is predicted to interact with E701 by a H-bond. **C**: Substitution Y705C is predicted to abolish the hydrogen bond, which may lead to a disruption of the fold of the protein in this region.

Molecular modelling of GlyT2 predicted that Y705 was located at the membrane interface with the extracellular space, with the side chain orientated inside the membrane towards E701, forming a hydrogen bond (Fig. 3.9). Substitution of the polar residue tyrosine with non-polar cysteine is likely to abolish this hydrogen bond, which may in turn disrupt the fold of the helix in this area. Furthermore, the Y705C substitution introduces a reactive thiol group in the place of a hydroxyl-containing side chain. Cysteines are frequently involved in disulphide bonding with other cysteines via the thiol groups. Thus, another plausible pathogenic mechanism for this mutation is the formation of disulphide bonds with other proteins, or between C705 and one of the 23 endogenous cysteines present in GlyT2 (Fig. 3.10), which could disrupt protein function or impair translocation to the membrane.

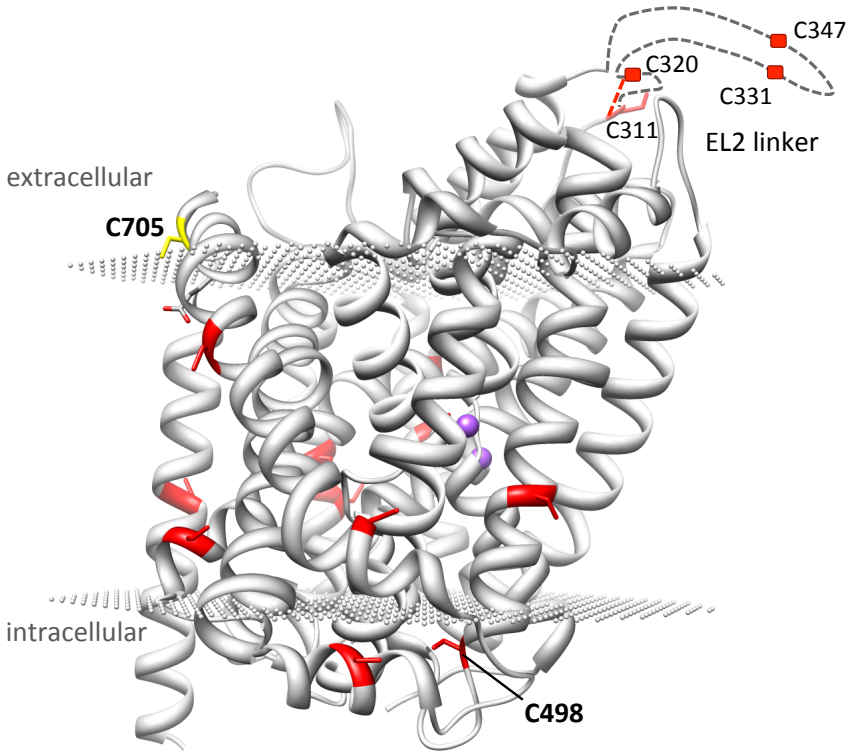

Figure 3.10. Molecular model of GlyT2 showing the location of cysteines. Cysteine residues of human GlyT2 are coloured in red, and C705 in yellow. C311 and C320 located in the second external loop between TM3 and TM4 are predicted to form a disulphide bond.

The López-Corcuera group confirmed that C705 was accessible from the exterior of the plasma membrane by thiol labelling with impermeant cysteine methanethiosulphonate (MTSEA) reagents. They also demonstrated that membrane expression of Y705C was impaired, but could be restored to wild-type levels with the permeable thiol reagent DTT, suggesting that C705 is involved in aberrant disulphide bond formation (Giménez *et al* 2012). In cell-surface labelling experiments, where free thiols of exposed cysteines were labelled with MTSEA-biotin, they showed that sulfhydryl (SH)-labelled GlyT2 Y705 membrane expression was around 2-3 fold higher than that of wild-type GlyT2, suggesting that the Y705C mutation increases cysteine availability at the cell surface. They showed that the Y705 mutant formed an additional disulphide bond compared to wild-type GlyT2. By using a pretreatment of N-ethylmaleimide (NEM), which modifies free thiols preventing SH-labelling, then adding DTT to cleave any disulphide bonds before finally labelling with MTSEA-biotin, they

found that the mutant retained more SH-label than wild-type. In order to identify the binding partner of Y705C, selected cysteine residues based on location whereby an intra- or intermolecular disulphide bond with C705 would be possible (Fig. 3.10), were individually substituted to serine to form double mutants. They found that C311S/Y705C or C320S/Y705C double mutants displayed a significantly reduced MTSEA-biotin labelling compared to the Y705C mutant, indicating that C705 can form disulphide bonds with either C311 or C320 (Giménez *et al* 2012). Although this region of GlyT2 cannot currently be modelled with any accuracy due to low sequence identity with LeuT, cysteines C311 and C320 are both exposed (located on the EL2 linker) and conserved with other members of the Na^+/Cl^--dependent neurotransmitter transporter family and have previously been shown to form a disulphide bond (Chen *et al* 1997; Chen *et al* 2007). Therefore, it is likely that the Y705C mutant disrupts GlyT2 cell surface trafficking and function by disrupting the C311-C320 disulphide bond and forming aberrant disulphide bonds with either C311 or C320.

3.8. Discussion

The studies described here have revealed the existence of several new pathogenic mechanisms that result in defective glycine transport. Previously known mechanisms included truncation of the GlyT2 polypeptide due to nonsense and frameshift mutations (Eulenburg *et al* 2006; Rees *et al* 2006), to which splice-site mutations can now be added. Truncating nonsense or frameshift mutations distributed throughout the GlyT2 coding region were found in 11/20 patients. As well as loss of function via protein truncation, these mutations also have the potential to lead to nonsense-mediated RNA decay. One nonsense mutation (R439X) was detected in several apparently unrelated cases from different ethnic backgrounds. Splice-site mutations were identified in cases 1 (IVS14+1ΔG), 7 (IVS13+1 G>T) and 9 (IVS8+1 G>A), all affecting the +1 position (in different introns). This means that the first nucleotide in the corresponding intron (i.e. part of the splice donor site consensus, 'gt') is deleted or substituted. These changes are predicted to result in either mis-splicing or intron retention.

Previous pathogenic mechanisms reported for missense mutations (Rees *et al* 2006) that resulted in apparently normal membrane trafficking included defects in glycine (W482R in TM6) or Na^+ binding (N509S in TM7). With the exception of S513I in TM7, the new missense substitutions described in this study (L237P, P243T, E248K, A275T, F547S, Y656H, G657A, Y705C) do not directly affect predicted Na^+ or glycine binding residues (Fig. 3.5). However, homology modelling of human GlyT2 based on the structure of the bacterial leucine transporter LeuT (Yamashita *et al* 2005) and analysis of key structure-function studies of related transporters has allowed us to suggest several hypothetical pathogenic mechanisms for hyperekplexia mutations. These include mutations that are predicted to: i) affect the conformation of TM2 (L237P, P243T); ii) indirectly affect glycine (A275T, E248K) or Na^+ (A275T, S513I) binding; iii) directly affect Cl^- binding (S513I); iv) affect the conformational mobility of EL4B (F547S); v) affect a cation-π interaction within or accessory protein interactions with the intracellular TM10-TM11 loop (Y656H, G657A); and vi) affect disulphide bonds within or between GlyT2 molecules (Y705C). I have also confirmed that the Cl^- binding residue S513 in GlyT2 (equivalent to S372 in SERT and S331 in GAT-1) has a key functional role in glycine uptake. Replacement of S372 in SERT with glutamate (S372E) or aspartate (S372D) results in Cl^- insensitive transporters with low uptake activity. However, consistent with these findings, substitution S372L completely abolished serotonin uptake (Forrest *et al* 2007) suggesting that either a Cl^- ion or a carboxylate side chain as in LeuT is required in this area for the normal uptake cycle.

The identification of these new GlyT2 mutations effectively triples the number of known individuals with presynaptic defects, firmly establishing mutations in the GlyT2 gene as a second major cause of startle disease (Carta *et al* 2012; Giménez *et al* 2012). Taken together with the recent study of mutations in the GlyR α1 subunit gene (*GLRA1*) (Chung *et al* 2010), it is now apparent that recessive hyperekplexia is far more common on a population basis than dominant hyperekplexia, which received initial attention due to the characterization of extended families (Shiang *et al* 1993; 1995) harbouring dominant *GLRA1* mutations such as R271Q/L. Genotype-phenotype

analyses have also revealed that *SLC6A5* mutations result in more severe clinical signs than mutations in *GLRA1*, including high rates of recurrent neonatal apnoea, learning difficulties and developmental delay (Carta *et al* 2012). One potential explanation for these differences is that defective presynaptic glycine uptake is predicted to affect the function of multiple GlyR subtypes, i.e. α2β, α3β as well as the predominant α1β isoform. In this context, it is noteworthy that GlyR α3 knockout mice were recently shown to have an irregular respiratory rhythm (Mantzke *et al* 2010).

4. CONGENITAL MUSCULAR DYSTONIA TYPES 1 AND 2 IN CATTLE FROM THE UNITED KINGDOM

4.1. Background

Livestock farming has changed dramatically over the last 50 years, due to the implementation of efficient breeding programs aimed at increasing 'performance indicators' such as feed conversion efficiency (a measure of the efficiency of converting feed mass into increased body mass), milk production and fertility. Whilst selection for these desirable traits increases profitability, it can also result in reduced effective population sizes due to inbreeding, especially in those species where artificial insemination is common. As a result of the practice of utilising so-called 'elite sires' with desirable traits, recessive genetic defects are beginning to emerge and cause significant economic and animal welfare issues. A classical example is bovine leukocyte adhesion deficiency (BLAD), a lethal immunodeficiency disorder affecting 14% of Holstein-Friesian bulls and 5.8% of cows at an estimated annual cost of $5 million (Shuster *et al* 1992). This disorder is caused by a missense mutation (D128G) in *ITGB2* encoding the leukocyte β2 integrin subunit CD18 (Shuster *et al* 1992). In a more recent study, 30% of Holstein-Friesian bulls were shown to be carriers of a missense mutation (V180F) in *SLC35A3*, encoding the Golgi UDP-N-acetylglucosamine transporter. This mutation causes complex vertebral malformation (CVM) resulting in premature birth, aborted foetuses and stillborn calves (Thomsen *et al* 2006). Curiously, study of cattle lineages has suggested that both of these genetic defects were spread throughout the Holstein-Friesian population by one founder sire: Carlin-M Ivanhoe Bell (Schuster *et al* 1992, Thomsen *et al* 2006). It is therefore in the interests of the breeding industry to develop methods for identifying and understanding the genetic causes of these recessive traits.

In 2008, my host laboratory was involved in a large-scale study aimed at understanding the basis for several newly emerging recessive disorders in cattle using genome-wide, high-density single nucleotide polymorphism (SNP) panels (Charlier *et al* 2008). This study reported the genetic basis of three inherited recessive disorders: congenital

muscular dystonia (CMD) types 1 and 2 in Belgian Blue cattle and ichthyosis foetalis in Italian Chianina cattle. Ichthyosis foetalis (IF) is a severe form of ichthyosis characterized by a thickening of the keratin layer in the foetal skin, which appears as large horny plates that are separated by deep cracks. Affected calves abort or die soon after birth (Molteni *et al* 2006). The genetic cause of IF was found to be a missense mutation (H1935) in *ABCA12* on bovine chromosome 2. This gene encodes an ATP-binding cassette (ABC) transporter found in the lamellar granules of keratinocytes, where it is thought to have an important role in the regulation of lipid trafficking (Akiyama *et al* 2006). Missense mutations in the first ATP-binding domain of human *ABCA12* cause lamellar ichthyosis type 2 (Lefévre *et al* 2003) whilst nonsense, frameshift and missense mutations underlie the more severe harlequin ichthyosis (Kelsell *et al* 2005; Thomas *et al* 2006).

Careful examination of CMD suggested two distinct phenotypes, classified as CMD1 and CMD2 (Vandeputte *et al* 2006). Calves affected by CMD1 have impaired swallowing, fatigue upon stimulation or exercise and muscle myotonia resulting in an inability to flex limbs and falling that could cause serious injury (Charlier *et al* 2008). CMD1 calves do not survive into adulthood but die within a few weeks of birth due to respiratory complications. By contrast, calves with CMD2 suffer severe episodes of muscle hypertonia upon acoustic or tactile stimulation and typically die within a few hours of birth. The gene for CMD1 was mapped to bovine chromosome 25 and DNA sequencing of candidate genes revealed a R559C missense mutation in the ATP binding site of the fast-twitch skeletal-muscle Ca^{2+} ATPase SERCA1, encoded by *ATP2A1* (Charlier *et al* 2008). SERCA1 is responsible for pumping Ca^{2+} from the cytosol back into the sarcoplasmic reticulum, thereby inducing muscle relaxation after contraction. Interestingly, a second *ATP2A1* missense mutation (R164H) was subsequently identified in congenital pseudomyotonia of Chianina cattle (Drögemüller *et al* 2008). Mutations in human *ATP2A1* cause Brody myopathy, a rare recessive disorder where affected individuals have exercise-induced muscle cramps and impaired muscle relaxation (Odermatt *et al* 1996; Vattemi *et al* 2010). By contrast, CMD2 mapped to

bovine chromosome 29 and was found to be a L270P missense mutation in *SLC6A5*, encoding the Na^+/Cl^--dependent glycine transporter GlyT2. In humans, missense, nonsense, frameshift and splice-site mutations in *SLC6A5* cause hyperekplexia (Rees *et al* 2006; Carta *et al* 2008), characterized by an exaggerated startle response to tactile or acoustic stimuli resulting in muscle hypertonia and life-threatening neonatal apnoea episodes. Introduction of the corresponding mutation (L269P) into the human GlyT2 cDNA by site-directed mutagenesis suggested that this mutation did not affect cell-surface trafficking of GlyT2, but abolished [^{3}H]glycine uptake mediated by recombinant GlyT2 in HEK293 cells (Charlier *et al* 2008).

This pioneering study opened the way for the characterization of further cattle disorders and a recent notable success was the characterisation of crooked tail syndrome (CTS) in Belgian Blue cattle (Fasquelle *et al* 2009). In this disorder, affected animals have a crooked tail and shortened head, growth retardation, extreme muscularity and spastic paresia - progressive stiffness and contraction (spasticity) in the lower limbs. Whilst CTS is not necessarily lethal, it does cause animal suffering and economic losses due to treatment of affected animals. CTS maps to bovine chromosome 19 and was recently found to be caused by different loss-of-function mutations in the bovine *MRC2* gene, encoding the mannose receptor C type 2, also known as 180 kDa Endocytic Transmembrane Glycoprotein (Endo180; Fasquelle *et al* 2009; Sartelet *et al* 2012). MRC2 plays a role in extracellular matrix degradation and remodelling, showing strong expression in developing bone. Whilst the link to muscular symptoms is less clear, a related mannose receptor is known to play a role in myoblast motility and muscle growth. In an initial study, a 2 bp deletion in *MRC2* was discovered in affected animals. The resulting frameshift reveals a premature stop codon that causes nonsense-mediated decay of the mutant messenger RNA. A second loss-of-function mutation - C636R, which causes defects in receptor oligomerization was also recently reported (Sartelet *et al* 2012). CTS is also peculiar in that the incidence of this disorder has risen very suddenly in the Belgian blue breed, with 25% of animals now estimated to be CTS carriers (Fasquelle *et al* 2009). One potential

explanation for this is that Belgian Blues are notorious for their exceptional muscular development known as 'double-muscling' (Figure 4.1). This extreme phenotype is related at least in part to an 11 bp loss-of-function deletion in the myostatin gene (*MSTN*) that has become fixed in the breed. Myostatin is a secreted growth differentiation factor that is a member of the TGF beta protein family that inhibits muscle differentiation and growth. Current theories suggest that *MRC2* mutations have been inadvertently selected by breeders as a result of their favourable effect on muscularity in animals that are heterozygous for *MRC2* mutations.

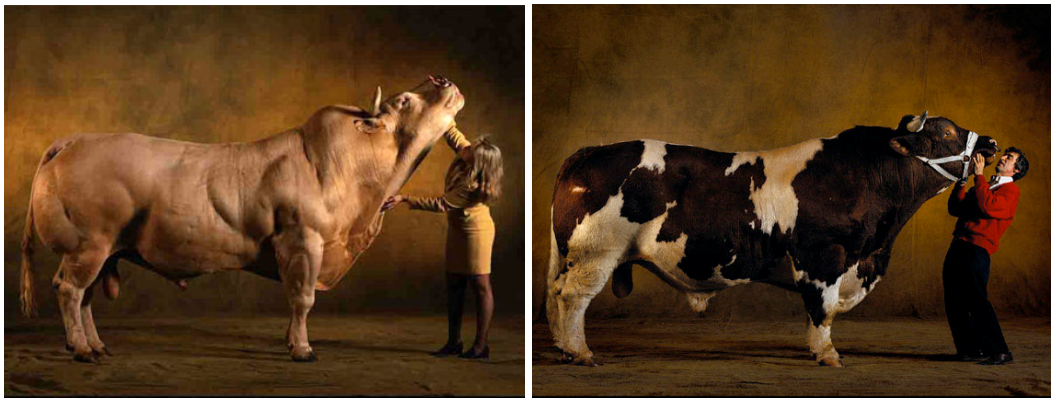

Figure 4.1. Images of Belgian blue sires showing characteristic 'double-muscling'. Double-muscling phenotype is an inherited condition, which results in the increased number of muscle fibres (hyperplasia) rather than the normal enlargement of individual muscle fibres (hypertrophy).

4.2. Study Aims

In 2009, two calves showing unusual neurological signs were presented to the Animal Health and Veterinary Laboratories Agency. Both calves exhibited what were to be recognized as classic clinical signs of congenital muscular dystonia 2 (CMD2). The aims of our study were:

- To investigate these two cases using molecular genetics approaches to determine whether a diagnosis of CMD2 could be confirmed and develop rapid genotyping tests for the L270P missense mutation.

- To raise awareness of genetic disorders in Belgian Blue cattle within the UK with the British Blue Cattle Society and to conduct a preliminary population analysis for CMD1, CMD2 and CTS in cattle from the United Kingdom.
- To use homology modelling to predict how the L270P mutation in bovine GlyT2 disrupts glycine transporter function.

Results

4.3. Detection of possible CMD2 cases in cattle from the United Kingdom

In November 2009, two calves showing unusual neurological signs were presented to the Animal Health and Veterinary Laboratories Agency (AHVLA). The calves were born on consecutive days to different dams by caesarean section and were the first offspring of a Belgian Blue bull. This bull and the two dams showed no clinical abnormality. By contrast, both affected calves exhibited persistent lateral recumbency and an unusually low head carriage, as well as transient extensor spasms following tactile or auditory stimulation (Figure 4.2a). These findings were present at birth, but they were non-progressive and the calves were alert, able to suckle and were relaxed between the episodes of spasm. The first calf died within 24 hours of birth, whilst the second was euthanized 48 hours after birth. Clinical examination of the live calf revealed abdominal breathing and slight hyperpnoea (increased depth of breathing). Withdrawal and patellar reflexes were also present, but difficult to assess since performing such tests induced spasms. There was obvious neck stiffness and the carpal joints could not be flexed more than 30°. Performing these manipulations stimulated spasms and caused distress to the calf. Both calves were of the expected size and weight. Bacteriological cultures were unremarkable and there was no evidence of congenital Bovine Viral Diarrhoea (BVD) infection. Macroscopic examination of the brain revealed a persistent cavum septum pellucidum and brain swelling (tentorial notch on occipital cortices) in both calves. No abnormality was detected on routine histological examination of brain (calf 1 and 2) or spinal cord (calf 1). These clinical signs suggested an inherited defect, since the observed muscle stiffness was not consistent with infectious conditions. Hypoxic causes and perinatal weak calf

syndrome were ruled out by clinical observations. The breed of cattle, the clinical signs and the lack of typical histopathological lesions were consistent with CMD2 (Charlier *et al* 2008).

4.4. Molecular genetic analysis of the bovine GlyT2 gene (*SLC6A5*)

Since the clinical signs of the two calves were consistent with CMD2, exon 4 from the bovine GlyT2 gene (*SLC6A5*) was amplified and sequenced in the two affected calves, the sire, both dams and nine additional animals from the affected farm (Table 4.1). Both calves were homozygous for the causative c.T809C mutation, providing confirmation of the diagnosis of CMD2 (Figure 4.2b). As expected, the sire and dams of the affected calves were heterozygous carriers of the c.T809C mutation (Figure 4.2b), as were three of seven animals in calf to the carrier sire (Table 4.1). This genotyping was confirmed by developing simple restriction fragment length polymorphisms (RFLPs; Figure 4.2c,d) using the enzymes *Fau*I and *Hpy*CH4V that detect the mutant (T809) and wild-type (C809) alleles, respectively. *Hpy*CH4V recognises the sequence **T**GCA and thus will only digest the wild-type GlyT2 allele at this position and not the mutant allele **C**GCA. On the other hand, *Fau*I will only cut the mutant allele at the site CC**C**GC, but not the wild-type allele CC**T**GC.

	Status	**Genotype**	**Outcome**
1	In calf to bull X	Normal	Calf normal
2	In calf to bull X	**Heterozygous (C/T)**	Calf normal
3	In calf to bull X	**Heterozygous (C/T)**	Calf normal
4	In calf to bull X	Normal (T/T)	Calf normal
5	In calf to bull X	**Heterozygous (C/T)**	Calf normal
6	In calf to bull X	Normal (T/T)	Calf normal
7	In calf to bull X	Normal (T/T)	Calf normal
8	Dam of bull X	Normal (T/T)	
9	Bull calf from dam of bull X	Normal (T/T)	
10	Bull X	**Heterozygous (C/T)**	
11	Dam of CMD2 calf 1	**Heterozygous (C/T)**	
12	Dam of CMD2 calf 2	**Heterozygous (C/T)**	
13	CMD2 calf 1	**Homozygous (C/C)**	
14	CMD2 calf 2	**Homozygous (C/C)**	

Table 4.1. Data from diagnostic DNA sequencing of *SLC6A5* exon 4 in fourteen cases. Note that three of seven animals in calf to the carrier sire were carriers, meaning that they had a 25% chance of producing an affected calf. Since all offspring had a normal phenotype at birth, these offspring have either a 67% (carrier dam) or 50% chance (non-carrier dam) of being CMD2 carriers.

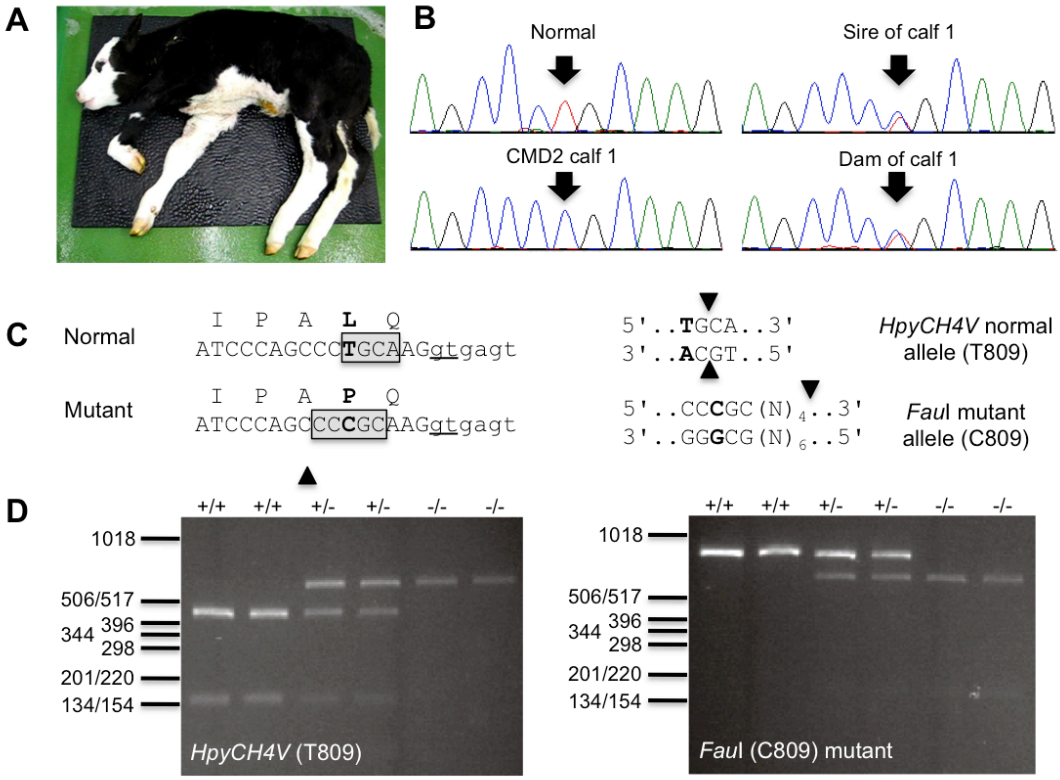

Figure 4.2. Genotyping confirms the identification of CMD2 cases in Belgian Blue cattle from the UK. A: One of two calves with CMD2 demonstrating lateral recumbency. **B**: Sanger DNA sequencing showing the homozygous T809C mutation in CMD2 calf 1, and the carrier status of the sire and dam. Coloured peaks represent the bases detected in the sequence. Two overlapping peaks at the same position represent a heterozygous base change. **C**: DNA sequences and protein translation of normal and mutant alleles showing the positions of the *Fau*I and *Hpy*CH4V restriction enzyme recognition sites used in RFLP analysis. **D**: RFLP analysis of duplicate samples demonstrating that *Hpy*CH4V detects the wild-type T809 allele, whilst *Fau*I digestion detects the C809 mutant allele in carriers and homozygous mutants. Banding patterns: *Hpy*CH4V: wild-type 475, 171, 69, 45, 27 and 24 bp. Mutant 646, 69, 45, 27 and 24 bp. *Fau*I: wild-type 811 bp, mutant 635 and 176 bp; The major products visible in the RFLP analysis are indicated by underlining.

4.5. Rapid genotyping assays for CMD2 – LightScanner high-resolution melting

Since approximately 55,000 Belgian Blue cattle are registered with the British Cattle Movement Service (www.bcms.gov.uk), screening for genetic mutations requires high-throughput methods. Since DNA sequencing of PCR amplicons is a relatively expensive method of genotyping, I decided to explore an alternative method of detecting the CMD2 mutation - LightScanner high resolution melting (HRM) analysis (Wittwer 2009). This method relies upon differential PCR amplicon melting in the presence of LCGreen Plus Dye (Figure 4.3). A variant of this method was also employed that used a LunaProbe (Montgomery *et al* 2007) a single 3'-blocked, unlabelled oligonucleotide probe that can be used to generate melting curves characteristic of the genotype under the probe (Fig 4.4c).

To test this high-throughput technique, I confirmed the genotypes found in our index cases (Figure 4.4d) and carried out further genotyping of fourteen unrelated Belgian Blue sire DNAs that were kindly provided by Pam Weiner (Roslin Institute). Surprisingly, two out of fourteen of these sires were carriers of the CMD2 mutation. From the British Blue Society database (http://www.britishbluecattle.org/database), it was estimated that the three carrier sires identified to date had generated a total of 1017 offspring and 1432 second-generation offspring. The decision was therefore taken to carry out a wider population analysis in collaboration with the British Blue Cattle Society and John Woolliams at the Roslin Institute, University of Edinburgh.

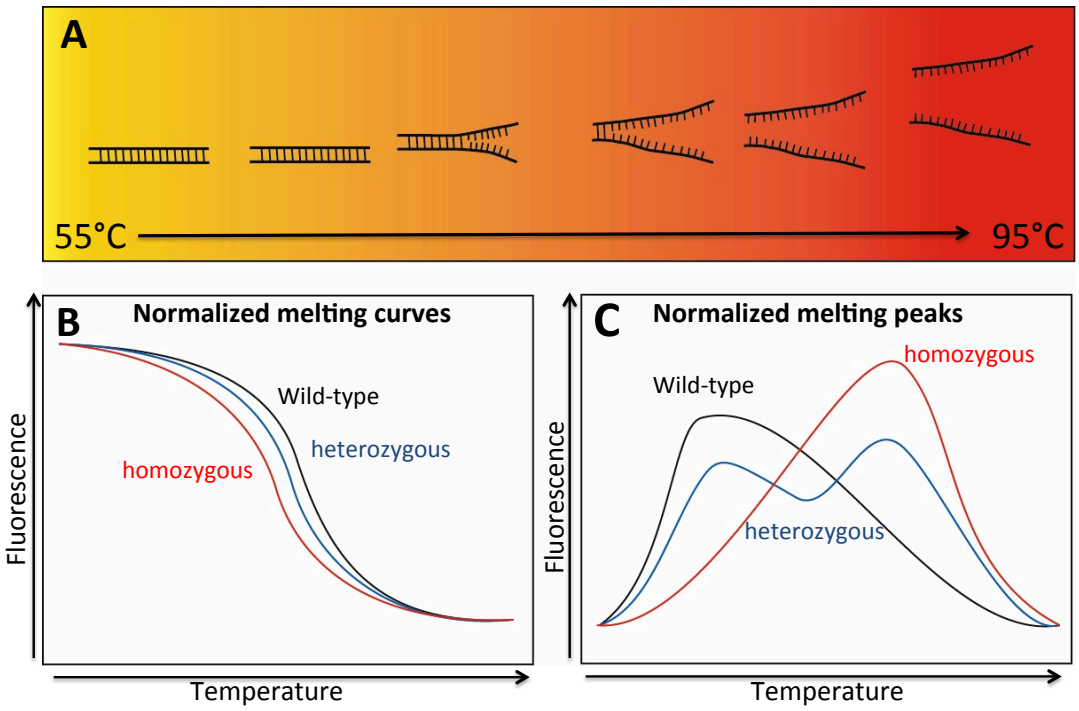

Figure 4.3. A schematic representation of high resolution melting (HRM). A: An illustration to show how a DNA double helix separates or "melts" apart as temperature increases. The temperature at which an amplicon melts is directly related to its sequence, and even a single base difference between two sequences will lead to different melting temperatures. HRM involves the addition of intercalating dyes that bind dsDNA and fluoresce, and then as the stands melt apart they become unbound. This process is monitored in real time by recording the amount of fluorescence as temperature increases then plotting to generate a melt curve. **B, C**: DNA that is homozygous for a mutation (red curves) will melt at a different temperature to the wild-type (black curves). DNA that is heterozygous for the same mutation will also have a distinct melt curve since half of the DNA will melt at the same temperature as the wild-type and half will melt at the same temperature as the mutant (blue curves).

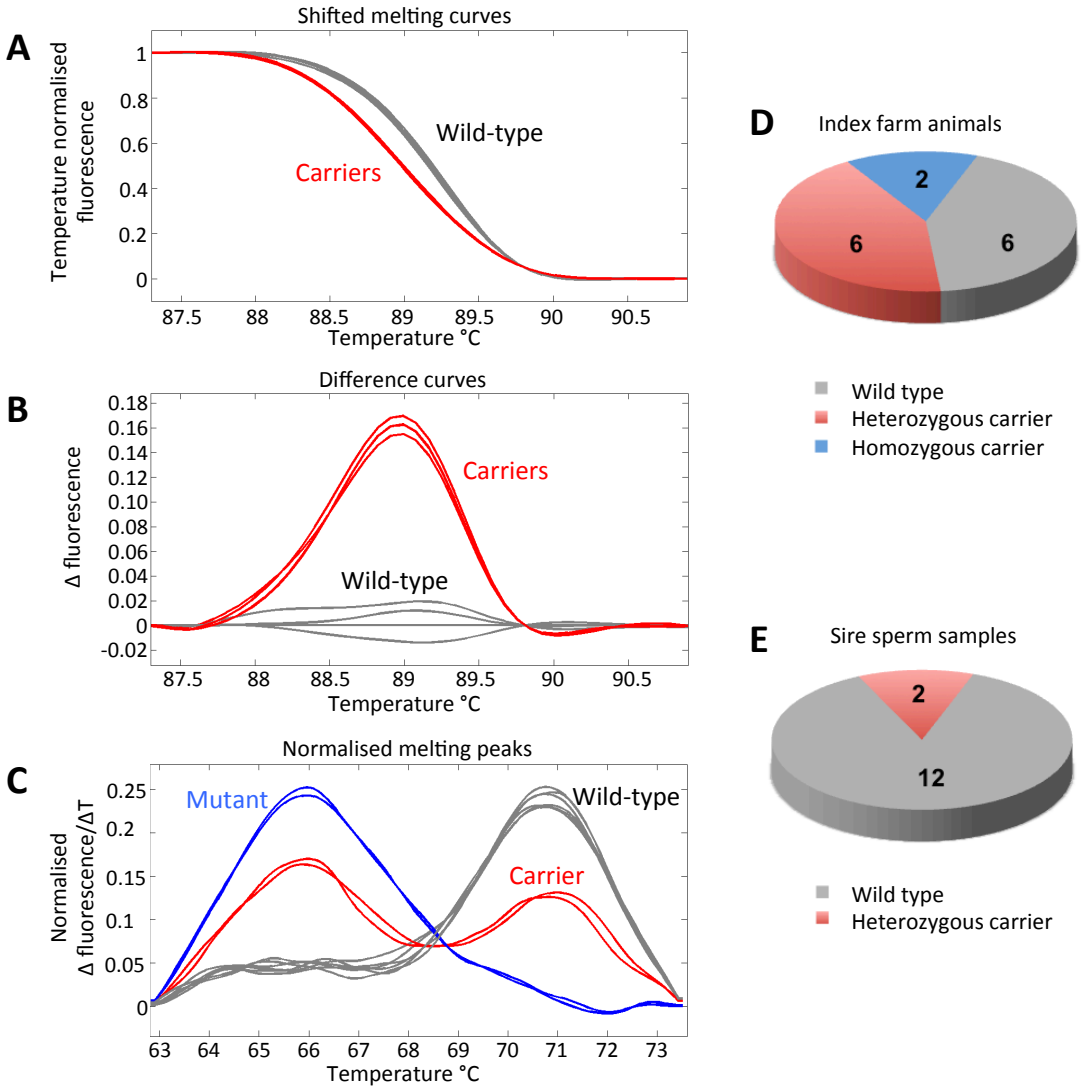

Figure 4.4. LightScanner HRM detection of CMD2 carriers. A,B: LightScanner melting curves and difference curves showing HRM analysis of a selection of *SLC6A5* exon 4 amplicons. Note that PCR products from heterozygous carriers (red curves) melt at lower temperatures than wild-type amplicons (grey curves) because of the mismatch caused by heterozygosity. **C**: Genotyping using asymmetric PCR and an unlabelled LunaProbe that spans the SNP of interest is complementary to the overproduced strand and has been blocked at the 3' end to prevent extension. As the product is melted, the LunaProbe melt allows detection of the mutant c.809C allele (blue curves), the heterozygous state (red curves) and the wild-type c.809T allele (grey curves). **D,E**: Detection rates for CMD2 homozygotes, carriers and wild-type animals in index farm animals and a sample of Belgian Blue sires.

4.6. Population analysis of CMD1, CMD2 and CTS

To estimate the frequencies of the deleterious variants causing CMD1, CMD2 and CTS, a sample of 49 bulls and a sample of 59 cows were genotyped at the appropriate loci. The bulls sampled were all those considered to be current contributors to the British Blue Cattle Society (BBCS) of which almost all were being used for artificial insemination. Of the 49 bulls, DNA was obtained from semen for 46 bulls, and from nasal swabs (DNA Genotek Inc., Kanata, Canada) for 3 bulls. The sampling design for cows was intended to span at least four age groups in order to obtain a clear picture across cohorts within a generation. All were sampled using nasal swabs. The BBCS requested samples from 100 cows and distributed swabs accordingly, however only 59 samples were returned to Roslin. One semen straw was duplicated (sexed/unsexed) making a total of 109 DNA samples processed. DNA was extracted at Ark Genomics at the Roslin Institute. The quality of the extracted DNA was assessed for suitability for Illumina SNP genotyping at KBioscience (Hoddesden, UK; www.kbioscience.co.uk). All 109 samples were genotyped for three loci: *ATPA2A1*, *SLC6A5* and *MRC2* which underlie the diseases CMD1, CMD2 and CTS respectively. Of the 63 samples obtained with nasal swabs, 10/63 (16%) failed. Note, again that failures cannot be assumed to be due to the capacity of the swab if used correctly, but are likely to reflect the adequacy of the sampling by the farmers. Our contribution to this research effort was to re-test all genotyping failures, mutation carriers and a sample of non-carriers at these loci. Confirmatory genotyping was performed using standard Sanger sequencing of PCR products and in addition an RFLP was developed for CMD1, based on digestion with *Dra*III, which recognises the sequence CACNNNGTG. There is a gain of *Dra*III recognition site in the CMD1 mutant allele. The wild-type DNA would only be digested in one position leading to two fragments of 39bp and 353bp. DNA of an affected animal would be digested in two positions leading to three fragments of 39bp, 176bp and 177bp. Digestion of DNA from a carrier would therefore lead to a total of four fragments (39bp, 176bp, 177bp and 353bp) since 50% of the DNA present would be digested in one position as is the case for wild-type, and the other 50% would be digested in three positions.

CMD1	Bulls	Cows	Total
Non-Carrier	51	54	105
Carrier	0	**3**	**3**
Total	51	57	108

CMD2	Bulls	Cows	Total
Non-Carrier	48	53	101
Carrier	**3**	**1**	**4**
Total	51	54	105

CTS	Bulls	Cows	Total
Non-Carrier	47	49	96
Carrier	**4**	0	**4**
Total	51	49	100

Table 4.2. Frequencies of carriers and non-carriers of CMD1, CMD2 and CTS in British Blue bulls and cows.

Collectively 6/51 bulls (12%) and 4/57 (7%) of cows were identified as carriers of at least one deleterious variant - one bull carried two variants. Thus, the carrier frequency for these disorders in British Blue population is low. For CTS there was a possible indication of an association with selection objectives since carriers were more frequent in the more selected bulls than in the randomly sampled cows. Regarding our validation, for CMD1, I obtained genotypes for six samples that failed with KBioscience SNP genotyping and also corrected one genotype – initially typed as a homozygous mutant by KBioscience, but was confirmed to be a carrier by Sanger sequencing. For CMD2, there were no disagreements with KBioscience SNP genotyping.

4.7. Molecular modelling of bovine GlyT2 suggests a mechanism for the CMD2 mutation

In this aspect of the project, I aimed to determine how the L270P mutation disrupts GlyT2 function. The L270 residue is not predicted to be directly involved in glycine or Na^+ binding (Charlier *et al* 2008; Figure 4.5 upper panel), however, a leucine residue is conserved at this position, as is the entire helix on which it resides, in a number of vertebrate GlyT2 sequences (Figure 4.5 lower panel), suggesting that this residue has a particular structural and/or functional importance.

Figure 4.5. Bovine GlyT2 sequence and conservation of L270. Upper panel: Sequence of bovine GlyT2 protein highlighting the 12 membrane spanning domains (TM1-12), the predicted Na^+ (blue triangles) and glycine binding residues (black circles) and the L270P mutation (black shading). Lower panel: Alignment of TM3 sequences from different vertebrate species showing the remarkable conservation of this membrane-spanning domain and the L270 residue.

The introduction of a proline at residue 270 can provide significant local structural rigidity to the three-dimensional structure of proteins, due to the rigid nature of the residue. Therefore, introduction of a proline in this position, at the turn of TM3 (shown in Fig. 4.9 and 4.10) could possibly destabilise the native fold, resulting in accelerated degradation or aggregation of the mutant protein (Bajaj *et al* 2007; Hurst *et al* 2009).

Fold recognition of bovine GlyT2 was performed with GenTHREADER (Lobley *et al* 2009) and HHPRED (Söding *et al* 2005) in order to identify the closest homologue available in the PDB. The bacterial leucine transporter (LeuT) (PDB: 2A65; Yamashita *et al* 2005) was identified as the best template structure (GenTHREADER p-value: 2e-21; HHPred e-value: 0). A profile-profile alignment between the bovine GlyT2 sequence and the LeuT structure was generated using the MUSCLE (Edgar 2004) and T-COFFEE (Notredame *et al* 2000) web servers, ultimately resulting in 26% sequence identity. Based on this alignment (Fig. 4.6), 50 homology models of bovine GlyT2 were calculated using MODELLER-9v7 (Eswar *et al* 2008), including heteroatoms (two Na^{2+} ions in the glycine binding site) and aligning the bound leucine with glycine. A large part of the extracellular loop between TM3 and TM4 (residues 313-355) was removed since it has little or no sequence identity with the template or any other known structure and therefore was not possible to model this region with any accuracy.

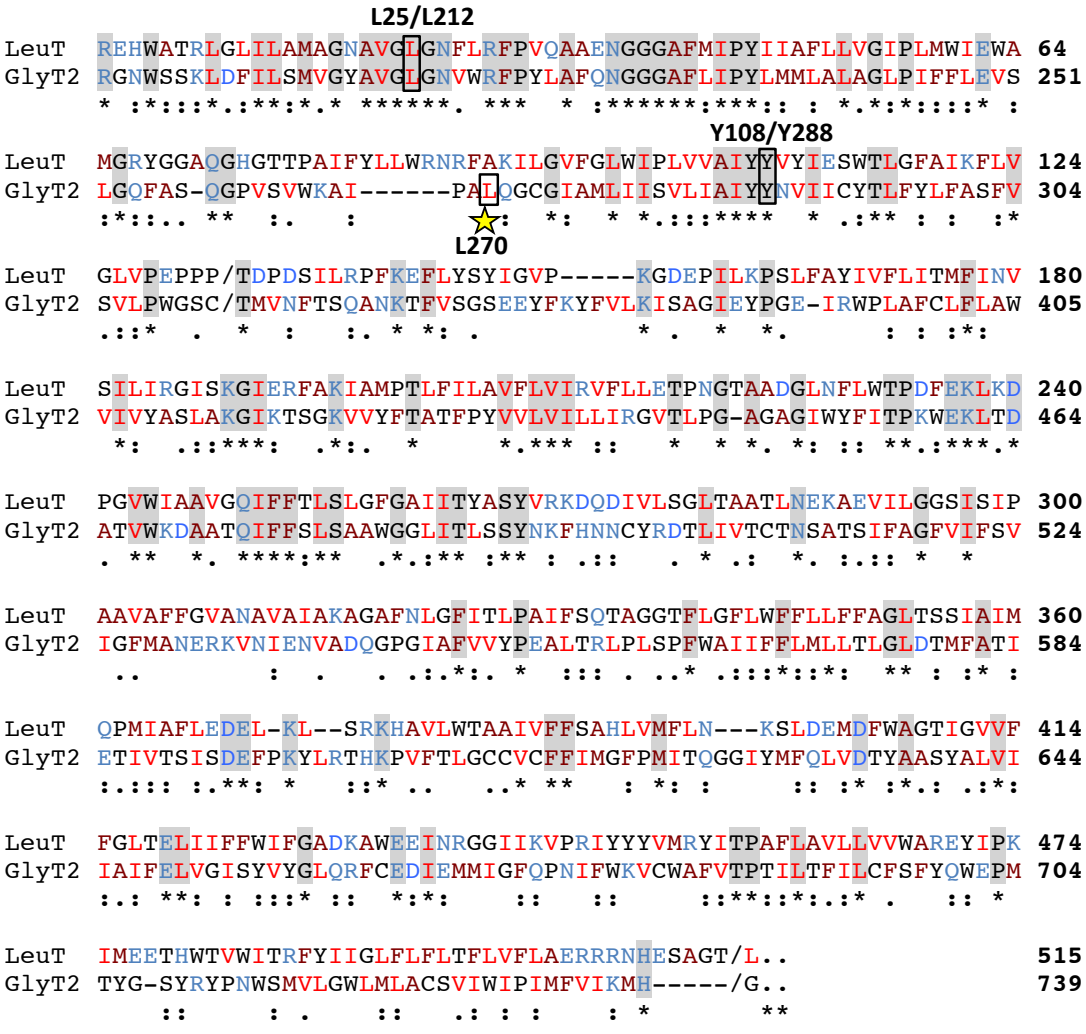

Figure 4.6. Alignment of LeuT and bovine GlyT2 sequences. Residues are coloured according to their hydrophobicity using the Kyte-Doolittle hydrophobicity scale (Kyte & Doolitte 1982), where the most hydrophobic residues are coloured in bright red (I, V, L), then less hydrophobic in dark red (F, C, A, M) and black (G, T, W, S, Y, P). Hydrophilic and charged residues are coloured in blue (D, E, N, Q, R, H, K). Residues highlighted in grey are conserved between LeuT and GlyT2 (also denoted by *) and residues that are partially conserved (grouped by similar residue characteristics, such as charge, aromaticity, polarity and size) are indicated with a colon or stop. Gaps are indicated with a hyphen (-) and a forward slash (/) indicates a chain break.

Each model was assessed with the DOPE statistical potential score (Shen & Sali 2006) and the best model based on the lowest score was selected. The GlyT2 homology model was calculated to have a normalized DOPE Z-score of -1.162, indicating a reasonable model, particularly considering the use of an alignment with less than 30%

sequence identity. The DOPE score of each residue of the selected model was plotted alongside the corresponding residues in the LeuT template structure (Fig. 4.7). Peaks of high energy, particularly DOPE scores that are more than 0, are considered unfavourable. Areas in the plot where the energies overlap are usually correlated with regions of high conservation between target and template. Generally, regions of particularly high energies were in agreement with those of LeuT, suggesting a native-like structure and a good model (Fig. 4.7).

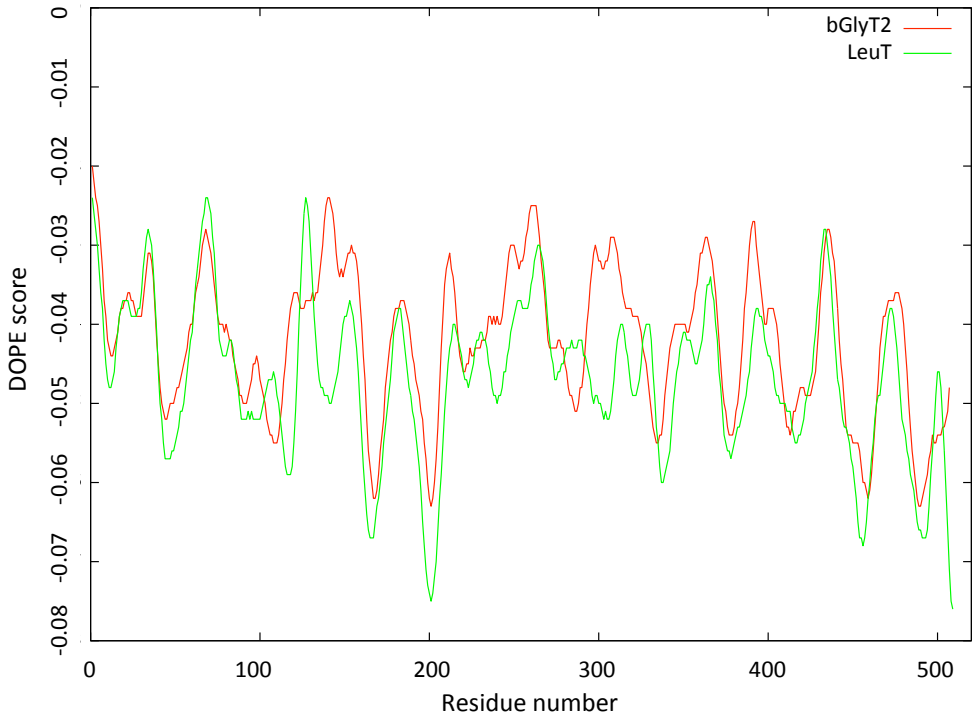

Figure 4.7. DOPE energy profiles of bovine GlyT2 homology model and LeuT structure. The DOPE score for each residue of both the GlyT2 homology model and the LeuT crystal structure plotted against the sequential residue number (1-500).

Evaluating this model using the QMEAN web server (Benkert *et al* 2009) indicated a reasonable model considering that it is a membrane protein (QMEAN Z-score = -3.8, which falls close to the range of the Z-scores typically found for native proteins of similar size) (Fig 4.8a,b). It is worth noting that the model is calculated to have very low residue error in the core region that surrounds the glycine-binding site (Fig. 4.8c).

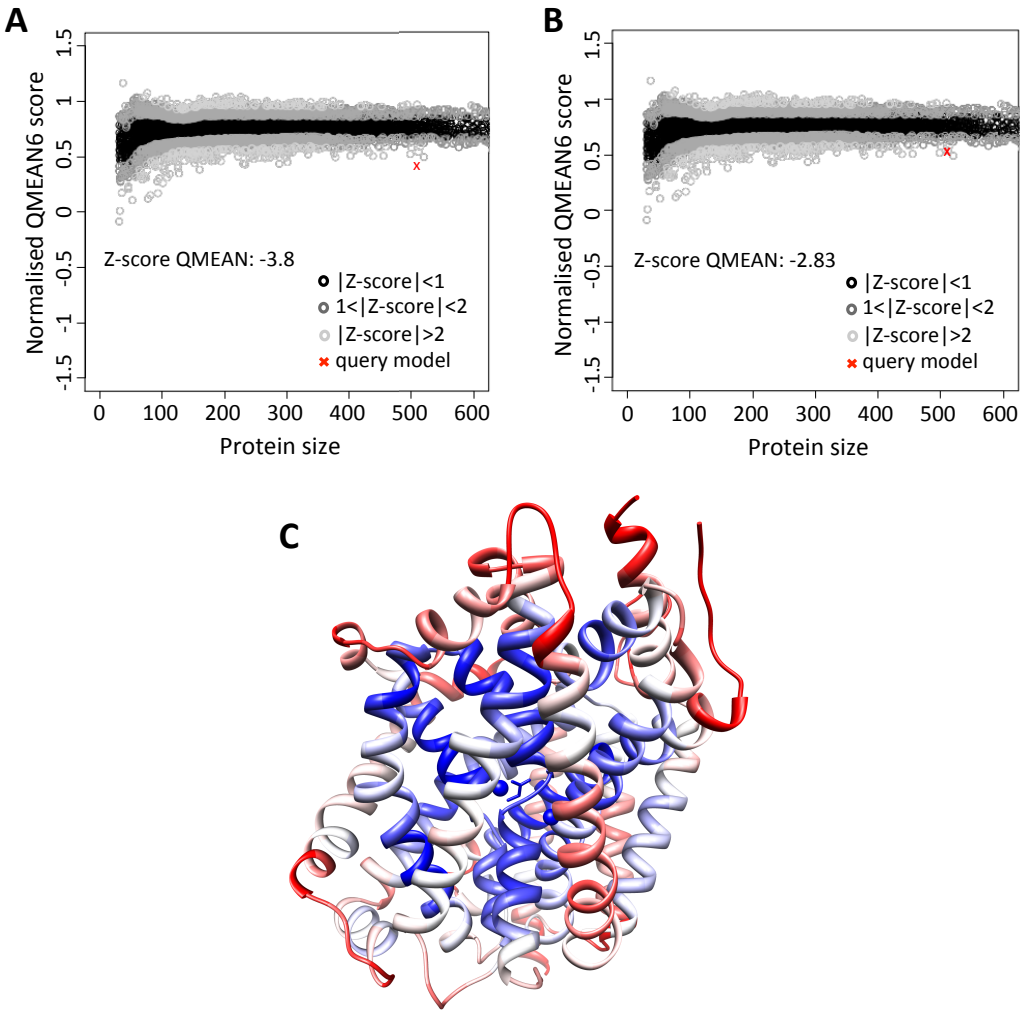

Figure 4.8. QMEAN scores of bovine GlyT2 homology model and LeuT structure plotted against scores for other native proteins. Normalised QMEAN6 scores for bovine GlyT2 homology model (**A**) and LeuT structure (**B**) plotted alongside a sample of other scores for crystal structures of other native proteins. The score for the GlyT2 model lies just outside of the cluster of scores for proteins of a similar size, and the score for the LeuT structure lies within the cluster but on the border. There is predicted to be low residue error in the core of the model (**C**), which is shown in blue. Areas of the model with higher predicted residue error are depicted in red.

Leu270 is predicted to be found on the protein exterior, and positioned at a turn in TM3 on the intracellular side of the membrane (Figure 4.9). The c.T809C mutation found in CMD2 cases, results in the native leucine at residue 270 being substituted for a proline. This residue is located in a short loop region, adjacent to another proline just two residues away at residue 268. It was not immediately obvious what the structural effects caused by this substitution in the proximate area might be. However, by using

information known about the leucine transporter, it was possible to make an informed hypothesis on how this mutation abolishes [^{3}H]glycine uptake (Charlier *et al* 2008). It is known from the crystal structure that the substrate and Na^+ binding sites of the LeuT are near the middle regions of TM3 and TM8 and alongside the unwound regions of TM1 and TM6 (Yamashita *et al* 2005). A tyrosine at residue 108 (Tyr108) located in TM3 of LeuT interacts with the carboxyl group of leucine substrate and also forms a hydrogen bond with the main-chain amide nitrogen of Leu25 located on TM1, which may stabilize the structure at the unwound region in TM1 (Yamashita *et al* 2005). Y108 corresponds to Y288 in the bovine GlyT2 model, and L25 to L212 (Fig 4.6). Substituting a proline in the place of leucine, just two residues down from another proline at the end of TM3 may cause a local distortion that could twist the helix, thus affecting the interactions further along TM3 (Figure 4.10). It is possible that this subtle distortion of the helix could move Y288 (Y108 in LeuT) so that the side chain can no longer interact with L212 on TM1 or the glycine substrate.

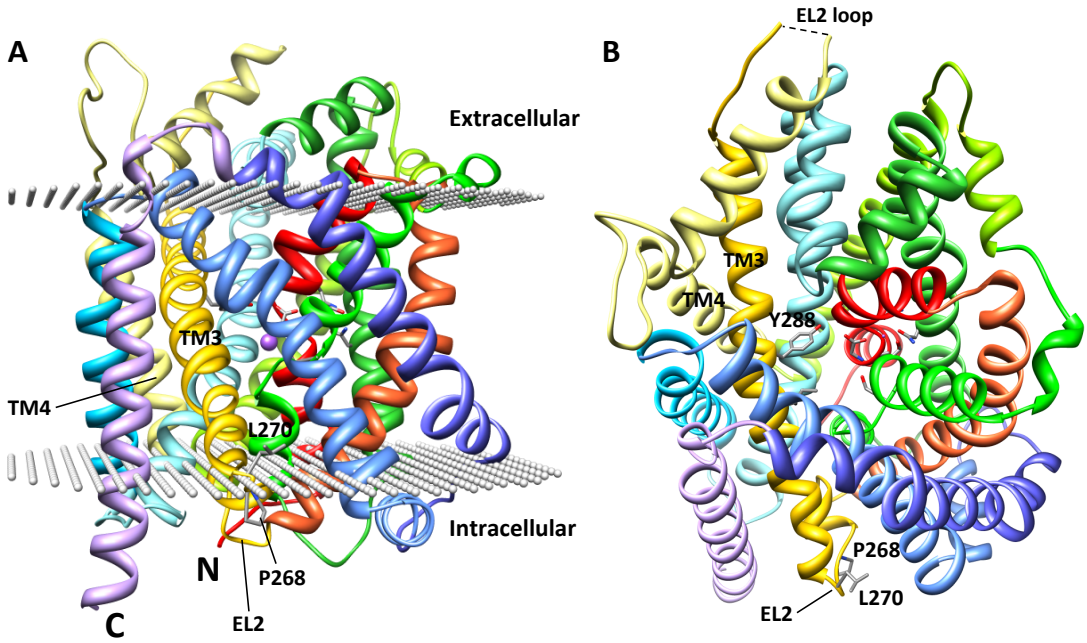

Figure 4.9. Molecular modelling of bovine GlyT2 suggests a molecular mechanism for the CMD2 mutation. **A, B**: Side and top views of the bovine GlyT2 monomer showing coloured transmembrane helices as ribbons (colours for each TM correspond to those in Fig 4.5). Glycine and two (out of three) sodium ions (purple spheres) are depicted. Note that the TM3-TM4 extracellular loop (EL2) was not modelled due to an insertion of residues 331-389 relative to LeuT.

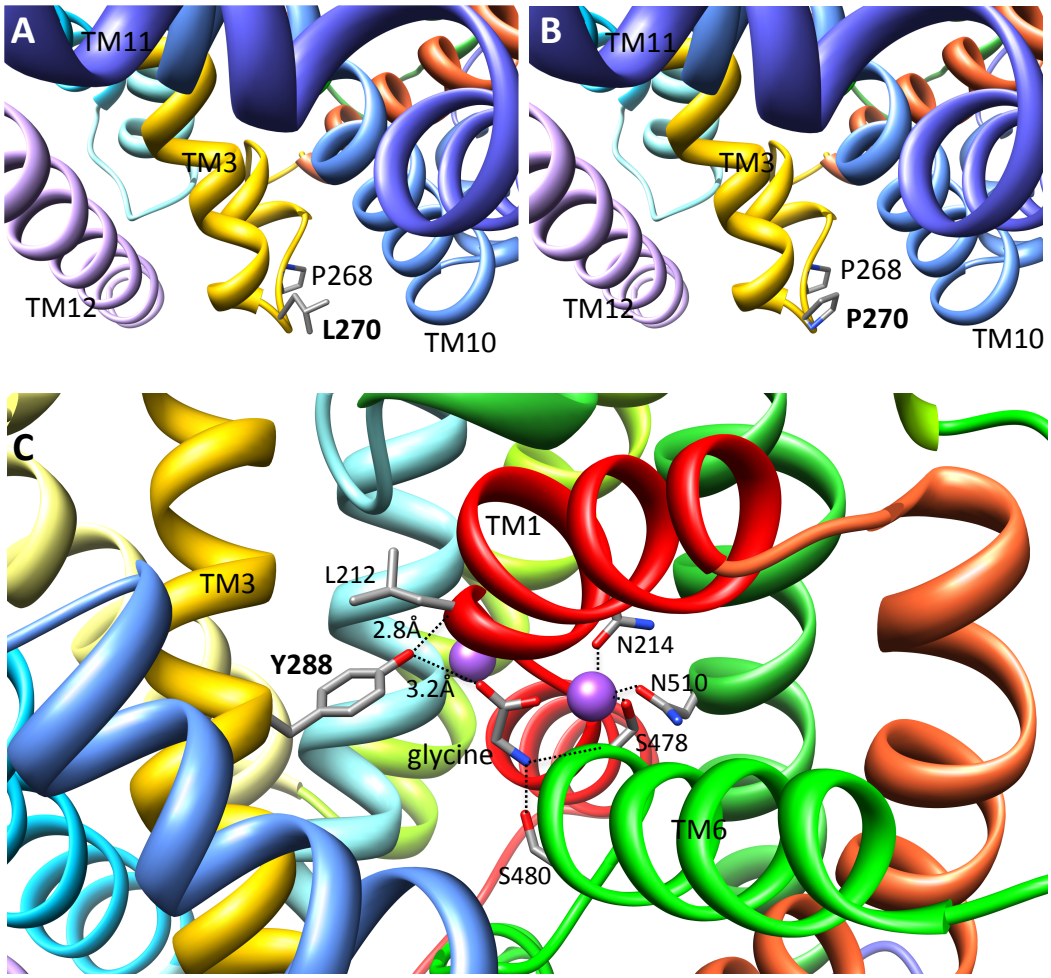

Figure 4.10. A suggested molecular mechanism for the CMD2 mutation. **A, B**: View of the positions of P268 and L270/P270 at the base of the TM3 helix and **C**: glycine and Na^+ binding residues, highlighting the contribution of Y288 in bovine GlyT2 to glycine binding and stabilization of the unwound region in TM1 via interaction with the L212 amide group.

4.8. Discussion

Two newborn Belgian Blue calves from a farm in the United Kingdom exhibited clinical signs consistent with congenital muscular dystonia type 2 (CMD2). DNA sequencing and RFLP analysis indicated that both calves were homozygous for the mutation c.T809C in *SLC6A5*, causing a L270P substitution in TM3 on bovine GlyT2. Further testing of animals from the index farm and a sample of Belgian Blue sires revealed an unexpectedly high frequency of CMD2 carriers. Fortunately, a second bull that was clear of the c.T809C mutation was identified (Table 4.1), ensuring a clear future breeding strategy. New inexpensive genotyping tests for the CMD2 allele were also developed that can be used to confirm diagnosis, identify carriers and guide future breeding strategy, so avoiding animal distress/premature death and minimizing the future economic impact of this disorder. I also took part in a collaborative study aimed at estimating the prevalence of CMD1, CMD2 and CTS in Belgian Blue cattle in the UK. For all three diseases, the population-wide frequency of carriers was predicted to be less than 5%. However, one limitation of the study was that it did not test for the new CTS allele recently discovered by Sartelet *et al* (2012). Since the Belgian Blue is of economic significance to the UK beef industry, I would make the following recommendations to the British Blue Cattle Society:

1. That the BBCS establishes a genetic screening program for all pedigree British Blue cattle.
2. Avoid the registration of any bull that carries any of the disease alleles.
3. Withdraw or limit the use of carrier bulls that remain active.

Using homology modelling of bovine GlyT2, based on the crystal structure of a bacterial leucine transporter, I also predicted that the L270P mutation is likely to disrupt GlyT2 function by altering the orientation of TM3, so disrupting interactions with L212 in TM1 or the glycine substrate. This sheds additional light on the potential pathogenic mechanism causing CMD2.

5. MOLECULAR ANALYSIS OF NOVEL GLYR β SUBUNIT MUTATIONS FOUND IN STARTLE DISEASE

5.1. Background

Inhibitory glycine receptors (GlyRs) are ligand-gated chloride channels enriched in the spinal cord, brainstem and retina, consisting of heteropentameric combinations of ligand-binding GlyR α subunits (α1-α4) together with the GlyR β subunit (Lynch 2009). Each GlyR subunit is composed of an N-terminal extracellular domain and four α-helical membrane-spanning domains (TM1-TM4), connected by short intracellular and extracellular loops. TM3 and TM4 are linked by a long intracellular domain containing binding sites for a variety of intracellular factors. Although GlyR α and β subunits both play an active role in agonist binding (Dutertre *et al* 2012; Grudzinska *et al* 2005), the GlyR β subunit was, until recently, widely assumed to play a purely structural role in heteromeric GlyRs. In part, this was due to the key role of the GlyR β subunit in mediating high-affinity interactions with the synaptic clustering molecule gephyrin (Meyer *et al* 1995), which in turn controls the dynamic localisation of GlyRs at synaptic sites (Feng *et al* 1998) More recently, the GlyR β subunit was also reported to interact with the proteins Vacuolar Protein Sorting 35 (Vps35) and Neurobeachin (Nbea), indicating a new role in GlyR trafficking (del Pino *et al* 2011). The TM2 domain of the GlyR β subunit also confers resistance to the effects of picrotoxinin (Pribilla *et al* 1992) and the insecticide lindane (Islam & Lynch 2012) and influences the main-state single-channel conductances of heteromeric αβ GlyRs (Bormann *et al* 1993).

Initial cross-linking studies of affinity-purified native GlyRs suggested that heteromeric GlyRs exist in a 3α1:2β subunit combination (Langosch *et al* 1988). However, recently three innovative studies have revealed that the GlyR β subunit represents the major component of heteromeric GlyRs, which are more likely to exist in a 2α1:3β stoichiometry with the subunit arrangement β-α-β-α-β. Grudzinska and colleagues (2005) compared the effects of mutations affecting predicted glycine binding residues in GlyR α1 and β subunits. Co-expression of mutant α1 with wild-type β subunits resulted in a full rescue of the GlyR EC_{50} value for glycine, whilst co-expression of wild-

type α1 with mutant β subunits resulted in a decrease in EC_{50}. A subsequent study (Dutertre *et al* 2012) revealed that the GlyR β subunit contributes more to agonist binding site formation than had been previously realised. Experiments with GlyR α1β concatemers also revealed that functional heteromeric GlyRs can be produced when co-expressed with GlyR β monomers but not when expressed alone, or with GlyR α1 monomers (Grudzinska *et al* 2005). This result was consistent with either a 2α:3β or a 1α:4β stoichiometry. However, quantification of radiolabeled methionine incorporation into recombinant α1β versus α1 GlyRs suggested a 2α1:3β stoichiometry and the subunit order β-α-β-α-β. This stoichiometry and subunit arrangement has recently been confirmed by imaging single antibody-bound GlyR α1β heteromers using atomic force microscopy (Yang *et al* 2012).

Defects in the adult GlyR isoform (α1β) also have an important a role in disease. Mutations in *GLRA1*, encoding the GlyR α1 subunit, are the major genetic cause of startle disease/hyperekplexia in humans (Shiang *et al* 1993, 1995; Chung *et al* 2010) and cause similar disorders in mice (Buckwalter *et al* 1994; Ryan *et al* 1994; Holland *et al* 2006) and Poll Hereford cattle (Pierce *et al* 2001). In humans, hyperekplexia affects newborn children and is characterized by exaggerated startle reflexes and hypertonia in response to sudden, unexpected auditory, tactile or visual stimuli. By contrast, *GLRB* mutations are less frequent, but recessive mutations have also been found in three families with hyperekplexia (Rees *et al* 2002; Al-Owain *et al* 2012; Lee *et al* 2012), in the mouse mutant *spastic* and the zebrafish *bandoneon* mutant (Hirata *et al* 2005). However, one conundrum is why so few *GLRB* mutations are found in startle disease relative to *GLRA1* or *SLC6A5*, which encodes the presynaptic glycine transporter GlyT2 (Rees *et al* 2006; Carta *et al* 2012; Giménez *et al* 2012).

5.2. Study aims

A collaborative study with the University of Leiden Medical Centre DNA revealed two new missense mutations in *GLRB* in individuals with hyperekplexia, resulting in L285R and W310C substitutions in the second (TM2) and third (TM3) membrane-spanning

domains, respectively. An additional missense mutation in *GLRB* was recently reported (Al-Owain *et al* 2012) resulting in an M177R substitution in the large N-terminal extracellular domain. The aims of my study were:

- To build a homology model of the human α1β GlyR based on the crystal structure of a related ligand-gated ion channel - the glutamate-gated Cl^- channel (GluCl) from *C. elegans*
- To use this homology model to predict how startle-disease associated substitutions in the human GlyR β subunit disrupt receptor structure and function.
- To use site-directed mutagenesis to introduce these mutations into a human GlyR β subunit expression construct in the vector pRK5 for use in functional assays.
- To correlate the results of fluorescence-based imaging and electrophysiological assays with molecular modelling predictions to determine the pathogenic mechanisms of these different mutations.

Results

5.3. Identification of mutations in the human GlyR β subunit gene (*GLRB*) in startle disease

Individuals with a clinical diagnosis of hyperekplexia (non-habituating startle response, positive nose tap test, history of neonatal/infantile hypertonicity) were ascertained by referral from paediatricians from international centres. Informed consent for participation in the study and publication of clinical data was obtained by the referring clinician. DNA was extracted from whole blood taken from 50 affected individuals and relatives using a Gentra Puregene DNA purification Kit (Gentra Systems, Minneapolis, USA). Primers recognizing the ten *GLRB* coding exons and flanking splice sites were designed using the LightScanner Primer Design Software package, version 1 (Idaho Technology Inc, Salt Lake City, Utah). To avoid allelic dropout, all primers were placed in intronic regions that were devoid of single-nucleotide polymorphisms (SNPs) as revealed by National Center for Biotechnology Information databases. Single nucleotide variants (SNVs) were assessed against those found in NCBI dbSNP Build 136 and release 20100804 from the 1000 genomes project (Appendix. tables 1-2). This analysis revealed two novel sequence variants in *GLRB* in two individuals (Fig. 5.1a,b).

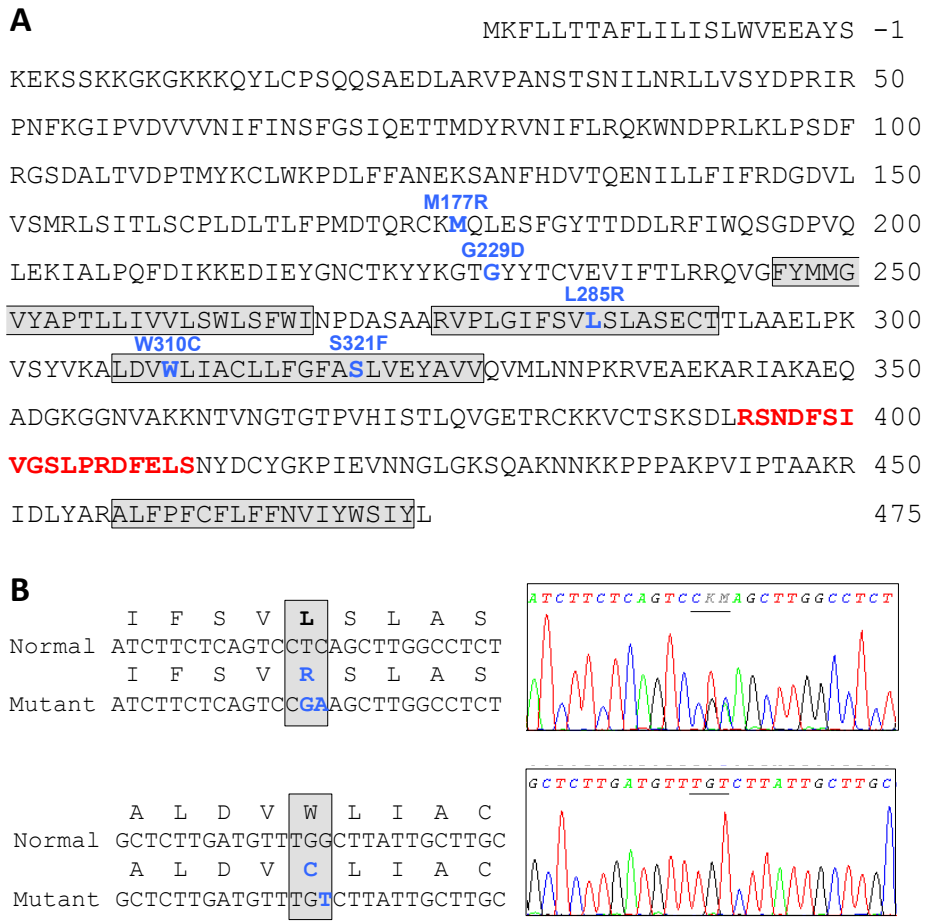

Figure 5.1. Human GlyR β subunit mutations identified in individuals with startle disease. A: Amino acid sequence of human GlyR β subunit (Uniprot: P48167) indicating the positions of putative transmembrane (TM) domains (grey shaded boxes) and amino acid residues affected by hyperekplexia mutations (Rees *et al* 2002; Al-Owain *et al* 2012; Lee *et al* 2012). Red text denotes the gephyrin-binding motif between TM3 and TM4. **B**: DNA sequence and electropherograms showing the c.920_921**Δ**insGA (p.L285R) and c.G996T (p.W310C) mutations identified in individuals 1 and 2.

Individual 1 harboured a two base missense mutation (c.920_921**Δ**insGA) resulting in a L285R substitution in the second membrane-spanning domain (TM2) of the GlyR β subunit. Since this change was found in the heterozygous state (Fig. 5.1b) and neither parent carries the mutation, this mutation appears to be *de novo*. By contrast, individual 2 harboured a different missense mutation (c.G996T), resulting in a W310C substitution in the third membrane-spanning domain (TM3) of the GlyR β subunit. At first glance, this mutation appeared to show recessive inheritance, since both parents

were heterozygous carriers. However, clinical assessments and functional data suggest an alternative mechanism of inheritance for this mutation (see below). The recessive mutation c.T596G, resulting in a M177R substitution, was recently reported (Al-Owain *et al* 2012) in nine individuals in a large family of Saudi Arabian origin with hyperekplexia and esotropia, an eye misalignment disorder where one or both eyes turn inward.

5.4. Clinical assessment of individuals harbouring novel *GLRB* mutations

Both infants, one male of white British origin (heterozygous for L285R) and one female of Turkish origin (homozygous for W310C), exhibited exaggerated startle and hypertonia in the neonatal period, requiring respiratory support. Neither was dysmorphic, but both were and remain small for their age. The male had transient dystonic posturing in infancy but then motor development largely normalized and startling became less prominent. Now nine years old, he plays football well and is making good progress in school and socially. Strabismus (eye misalignment) was observed at 14 months, but is no longer present. In this case, there was no family history of startle disease. The female developed hypsarrhythmia (abnormal EEG) in early childhood, which settled, but she is still treated with clonazepam, the typical pharmacological treatment for hyperekplexia. She does not have apparent visual or learning difficulties. Her older sister (untested) has clear symptoms and her younger brother (tested) is also affected. The parents (both heterozygous for W310C) are not currently symptomatic, but showed 'light symptoms' in early life. Both index cases had white matter changes on MRI, mild in the boy and periventricular cystic changes in the girl.

5.5. Building a homology model of the human GlyR $\alpha 1\beta$ isoform

M177R, L285R and W310C substitutions were predicted (Table 5.1) to be damaging using SIFT (http://sift.jcvi.org/; Ng 2003). Using Polyphen-2 (http://genetics.bwh.harvard.edu/pph2; Adzhubei *et al* 2010), the changes were predicted to be probably (M177R) or possibly damaging (L285R and W310C).

To gain further insights into the potential pathogenic mechanisms of these mutations, I constructed a homology model of the human α1β GlyR based on the crystal structure of the GluCl from *C. elegans* (PDB: 3RHW) (Hibbs & Gouaux 2011). The GluCl structure makes a significantly better template for homology modelling of the human GlyR compared to other ligand-gated ion channel structures such as *Gloeobacter violaceus* GLIC (Bocquet *et al* 2009), *Erwinia chrysanthemi* ELIC (Hilf & Dutzler 2008) or the *T. momorata* nAChR (Unwin 2005) due to functional similarity (i.e. GluCl is an anion channel) and higher overall sequence identity.

A profile-profile alignment between the human GlyR α1 and β subunits and the GluCl α subunit was generated using the MUSCLE and T-COFFEE web servers, which resulted in 44.4% (α1) and 38.9% (β) identity (Fig 5.2). By contrast, GlyR subunits share lower sequence identity with GLIC (α1: 21.8% and β: 24.6%) and different nAChR subunits (α1: 19.7 to 20% and β: 23.9 to 24.6%). A short part of the sequence was removed from the β subunit at the N-terminus (residues 1-30), and the extended loop region between TM3 and TM4 of both the α1 and β subunits were removed (residues 312-396 and 337-453, respectively) since it was not possible to model these regions due to very low sequence identity. The models were built with additional restraints enforcing disulphide bonds between cysteine residues involved in the cys-cys loops within each subunit (for GlyR α1: residues C138-C152 and C198-C209, for GlyR β: C161-C175 and C221-C233). Each model was assessed with the DOPE statistical potential score (Shen & Sali 2006) and the best model based on the lowest score was selected. The GlyT2 homology model was calculated to have a normalized DOPE Z-score of -0.222, indicating a reasonable model, particularly considering the use of an alignment with less than 50% sequence identity.

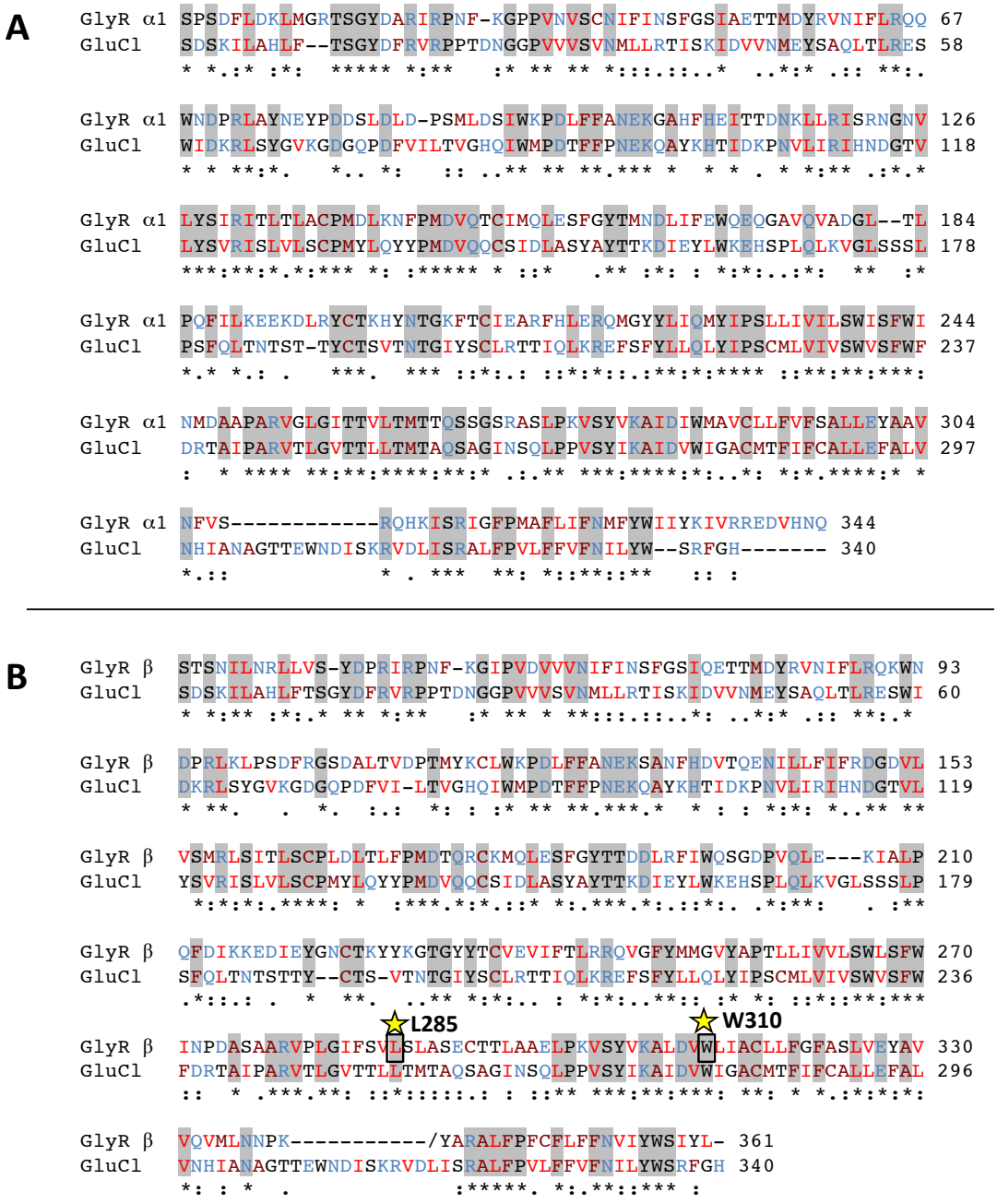

Figure 5.2. Alignment of GlyR and GluCl sequences. Residues are coloured according to their hydrophobicity using the Kyte-Doolittle hydrophobicity scale (Kyte & Doolitte 1982), where the most hydrophobic residues are coloured in bright red (I, V, L), then less hydrophobic in dark red (F, C, A, M) and black (G, T, W, S, Y, P). Hydrophilic and charged residues are coloured in blue (D, E, N, Q, R, H, K). Residues highlighted in grey are conserved between GlyR α1 or βb and GluCl (also denoted by a *) and residues that are partially conserved (grouped by similar residue characteristics, such as charge, aromaticity, polarity and size) are indicated with a colon or stop. Gaps are indicated with a hyphen (-).

The DOPE score of each residue of a single GlyR α1 subunit from the selected model was plotted alongside the corresponding residues in a single subunit from the GluCl template structure (Fig. 5.3a). The same was done for a single GlyR β subunit (Fig. 5.3b). Peaks of high energy, particularly DOPE scores that are more than 0, are considered unfavourable. Areas in the plot where the energies overlap are usually correlated with regions of high conservation between target and template. Generally, regions of particularly high energies were in agreement with those of LeuT, suggesting a native-like structure and a good model.

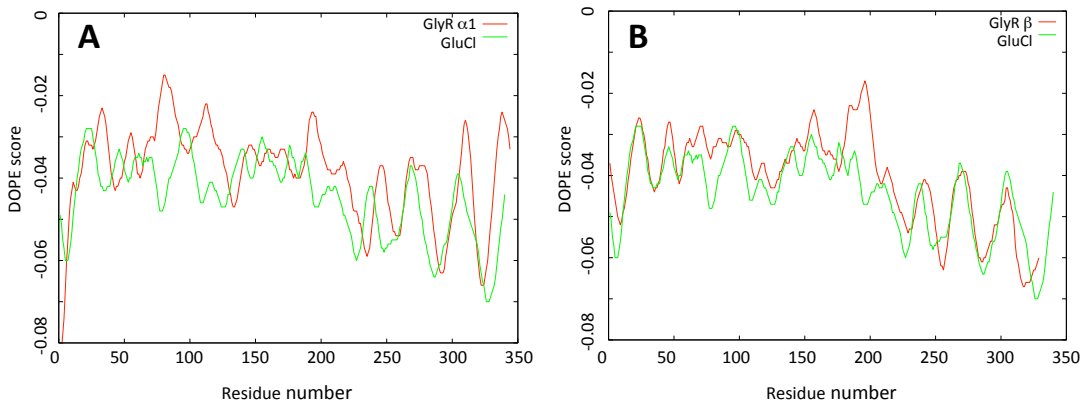

Figure 5.3. DOPE energy profiles of human GlyR α1 and β subunits compared with a single chain of GluCl. The DOPE score for each residue of both the GlyR α1 subunit (**A**) and the GlyR β subunit (**B**) from the GlyR α1β homology model, plotted against the sequential residue number alongside the DOPE scores for each residue of a single GluCl subunit.

Evaluating the model using the QMEAN web server (Benkert *et al* 2009) indicated a reasonable model considering that it is a membrane protein (QMEAN Z-score = -3.24), which falls close to the range of the Z-scores typically found for native proteins of similar size (Fig 5.4).

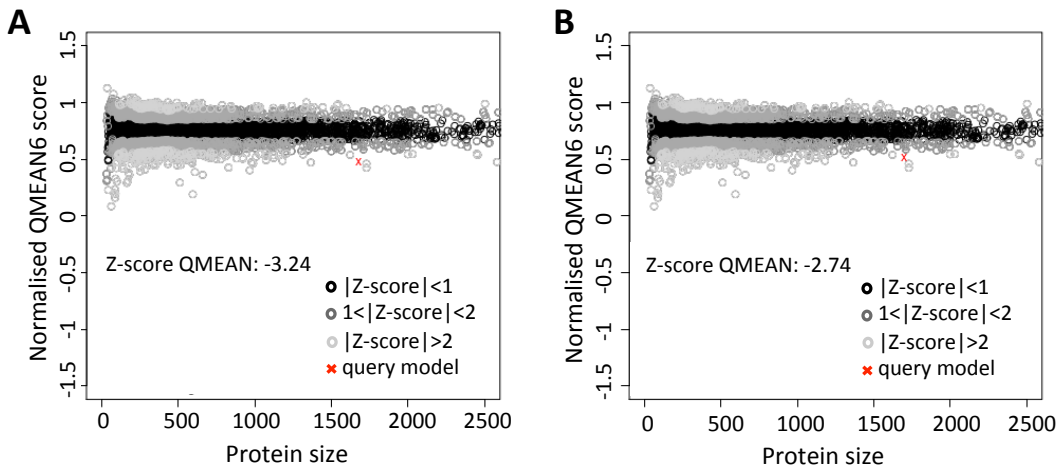

Figure 5.4. QMEAN scores of human GlyR α1β homology model and GluCl plotted against scores for other native proteins. Normalised QMEAN6 scores for human GlyR homology model (**A**) and GluCl (**B**) plotted beside a sample of other scores for crystal structures of other native proteins. The score for the GlyR model lies just outside of the cluster of scores for proteins of a similar size, as does the score for GluCl.

Selected non-synonymous substitutions were modelled into the GlyR homology model with the '*swapaa*' command in Chimera (Pettersen *et al* 2004) using the Dunbrack backbone-dependent rotamer library (Dunbrack 2002) and taking into account the lowest clash score, highest number of H-bonds and highest rotamer probability.

My GlyR model (Fig. 5.5a) predicts that L285 in the GlyR β subunit is a pore-lining residue, projecting a small, uncharged, hydrophobic side chain into the ion-channel lumen (Fig. 5.5b). The L285R mutation introduces a significantly larger, positively-charged side chain that projects further into the chloride channel (Fig. 5.5c). Because this mutation is *de novo* and the individual is heterozygous for L285R, in theory heteromeric GlyRs could be formed *in vivo* containing 3×wild-type, 1×, 2× or 3× mutant GlyR β subunits (Fig. 5.5c). Further modelling of pore-lining residues at a β subunit-β subunit interface (Fig. 5.6a,b) reveals that L285 is at the 9' position in TM2, equivalent to L254 in GluCl. This mutation is predicted to disrupt the pentameric radial symmetry of the 9' leucine 'hydrophobic girdle' that stabilizes the pore in the closed-channel state. Disruption of this girdle induces spontaneous channel activity and/or enhances agonist sensitivity in other Cys-loop receptors (Chang & Weiss 1998, 1999). In common

with GluCl, none of the major pore-lining residues are charged, with the exception of Asp-271 at the -5' position at the cytoplasmic end of TM2 (Fig. 5.6a,b). This is consistent with different models of anion-channel function which suggest that the positive electrostatic potential at the base of the pore arises from either: i) buried basic residues at the intracellular entrance of the pore (Cymes & Grosman 2011), such as R276 in the TM2 0' position in the GlyR β subunit (R252 in the GlyR α1 subunit) or ii) oriented peptide dipoles in the TM2 α helices (Hibbs & Gouaux 2011).

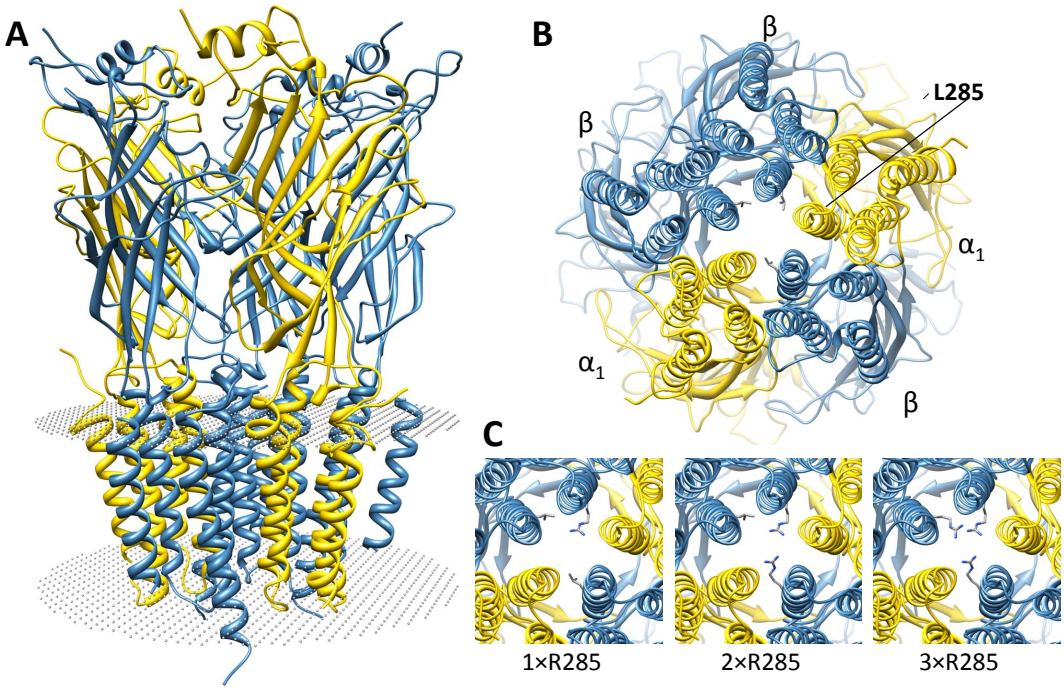

Figure 5.5. Molecular modelling of the human α1β GlyR and mutation L285R. A: Side view of the molecular model of the human GlyR α1β pentamer showing α1 subunits in gold and β subunits in blue. Grey spheres indicate the positions of the inner and outer membrane surfaces. **B**: Orthographic view of the GlyR α1β pentamer from the bottom, showing the relative position of the L285 side chain in the ion-channel pores. **C**: Panels depicting the multiple GlyR isoforms that are predicted to be formed on co-expression of wild-type and mutant GlyR β subunits harbouring the L285R substitution.

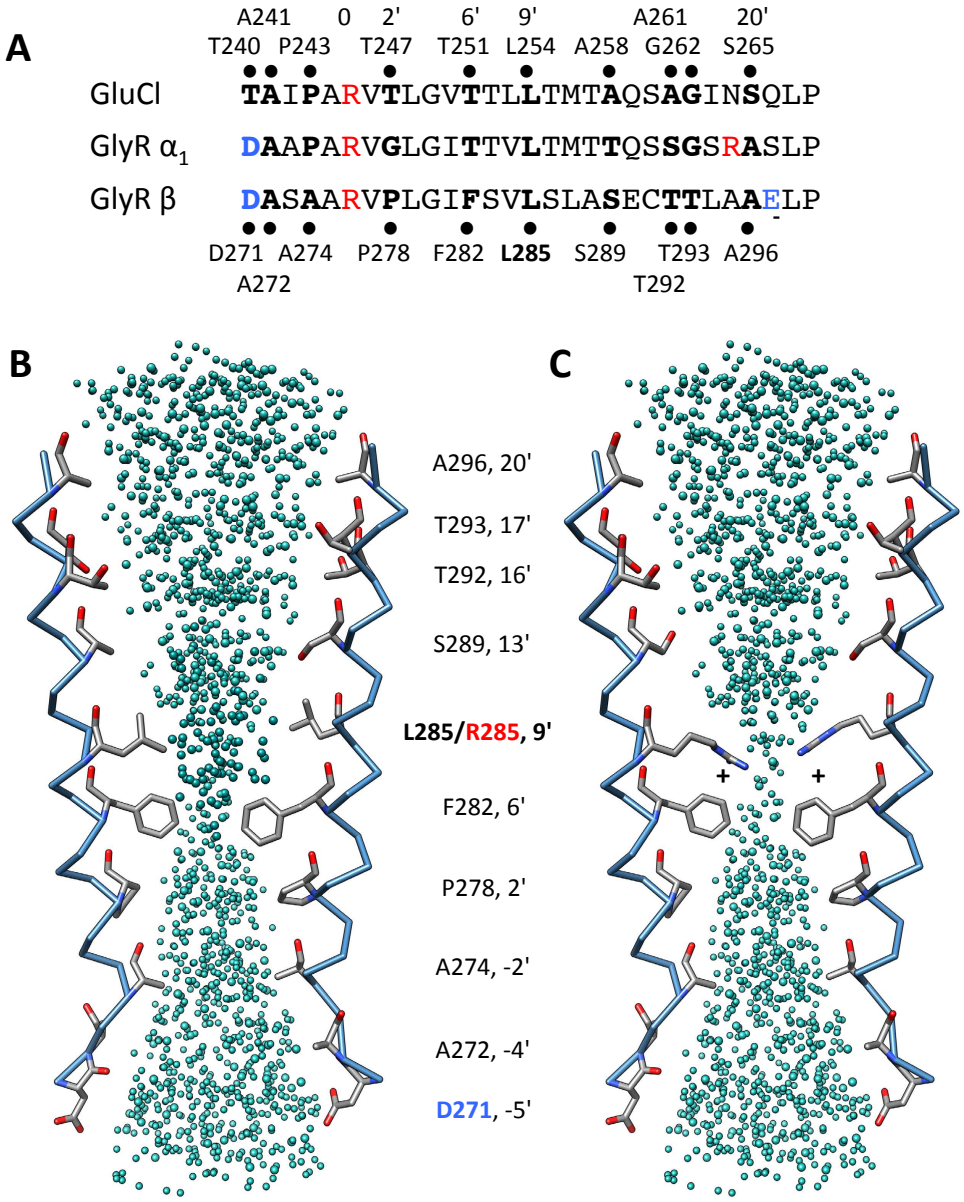

Figure 5.6. Mutation L285R inserts a charged residue into the GlyR ion-channel pore. **A**: Sequence alignment of TM2 of GluCl with GlyR α1 and β subunits. Relative positions of residues that line the Cl^- channel are indicated with black circles and residue numbers indicated for GluCl and the GlyR β subunit. **B, C**: Schematic diagrams depicting two GlyR β subunit TM2 α-helices and residues lining the Cl^- channel for wild-type L285 (**B**) and the R285 mutant (**C**). Blue spheres represent the predicted internal volume of the ion channel and side-chains are shown for pore-lining residues. The L285R substitution is predicted to create a premature narrowing of the pore at the 9' position and inserts a positively charged residue (+) into a helix that is uncharged, with the exception of D271 on the cytoplasmic face of the ion-channel pore.

By contrast, the side-chain of W310 in TM3 is part of an intramembrane network of aromatic residues (phenylalanine, tyrosine, histidine and tryptophan) contributed by TM1, TM3 and TM4 (Fig. 5.7a,b). Substitution W310C introduces a significantly smaller and shorter side chain into this position (Fig. 5.7d), with a reactive sulfhydryl group. Although I considered that C310 might form a disulphide bond with C314 in TM3, one turn down the helix in our model, the distance between the two sulphur atoms (4.6Å) precludes this interaction. In addition, disulphide bonds rarely form within the membrane or a reducing intracellular environment. Rather, my modelling predicts that the W310C substitution disrupts aromatic stacking (Fig. 5.7b) - a key requirement for correct GlyR assembly and cell-surface trafficking (Haeger *et al* 2009).

M177 is a hydrophobic amino acid located in the GlyR extracellular domain on a β-sheet (β7). The side-chain of this residue forms part of a hydrophobic pocket (Fig. 5.8a) that is conserved within the GluCl structure (Hibbs & Gouaux 2011), which has an isoleucine at the equivalent position. The M177R substitution results in the introduction of a positively-charged, hydrophilic side chain into this pocket (Fig. 5.8b), which is predicted to have a significant effect on the folding of the extracellular domain. Although M177/R177 are not predicted to directly influence glycine binding (Fig. 5.8c,d), disruption of the local structure around M177 is predicted to cause an indirect effect on agonist binding or signal transduction. For example, M177 is located just five residues away from F182, which is involved in a cation-π interaction (Pless *et al* 2011) with the amino group of the glycine agonist.

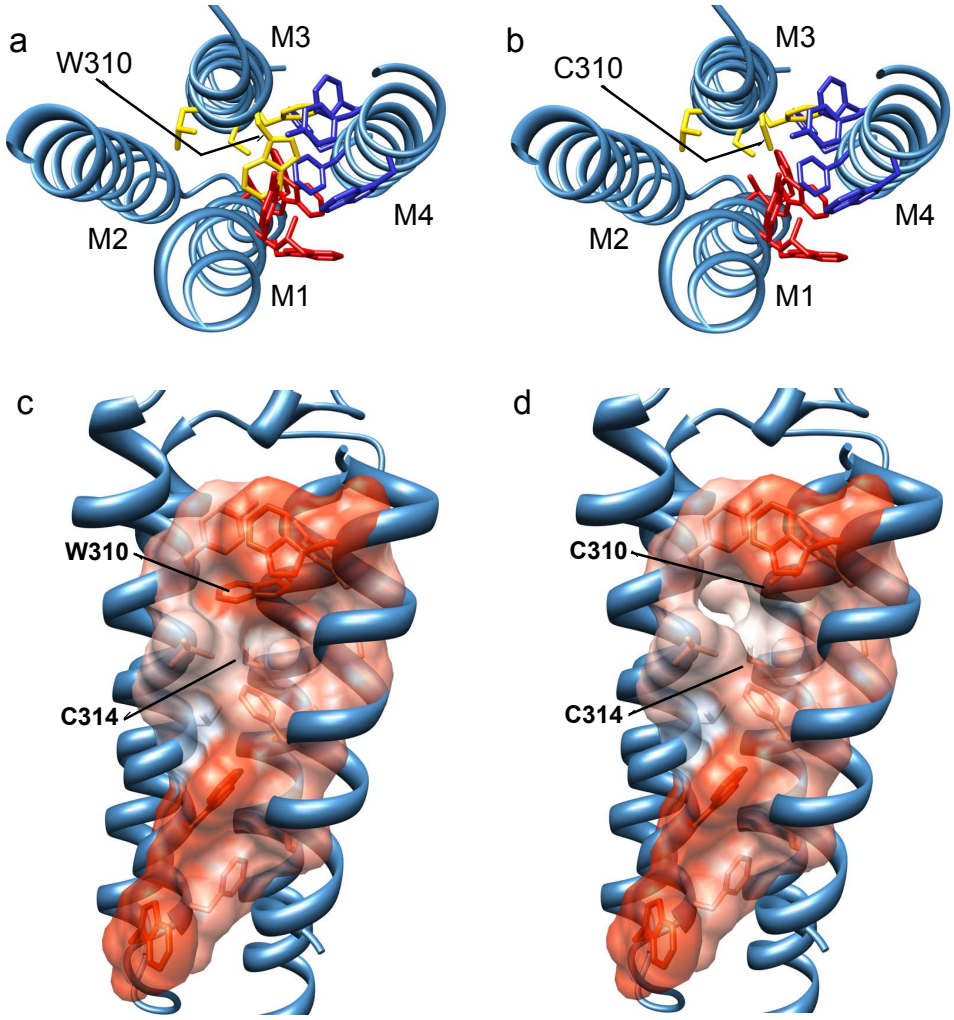

Figure 5.7. Mutation W310C disrupts hydrophobic stacking of membrane-spanning domains TM1-TM3. A: Cutaway view from the extracellular side of the GlyR β subunit showing how the side-chain of residue W310 in TM3 contributes to a hydrophobic stack of aromatic residues contributed by membrane-spanning domains TM1, TM3 and TM4. Hydrophobic side-chains of residues predicted to be involved in this stack are coloured according to membrane-spanning domain: TM1 – red, TM3 – yellow and TM4 – blue. **B**: Mutation W310C results the loss of the aromatic, hydrophobic side chain of tryptophan and replacement with the shorter, reactive cysteine sulphydryl side chain. **C,D**: Side view showing the hydrophobicity of individual residues in the stack by colouring the atom surfaces. Hydrophobic surfaces are depicted in orange, hydrophilic in blue and white indicates neutral polarity. The W310C substitution is predicted to disrupt the hydrophobic stack, altering the tertiary fold of the membrane-spanning domains. Disulphide bond formation is unlikely, since this is rare within the membrane or an intracellular environment and the only adjacent partner (C314 in TM3) is 4.6Å away from C310.

5.6. Novel GlyR β subunit mutations L285R and W310C impair GlyR function

To test these predictions, I used site-directed mutagenesis (QuikChange, Agilent) to introduce each mutation into a human GlyR β subunit cDNA cloned into the expression vector pRK5. These constructs were verified by sequencing and analysed by Anna Bode, Angelo Keramidas and Joseph Lynch at the Queensland Brain Institute, Brisbane, Australia. HEK293 cells were co-transfected with expression constructs encoding the GlyR α1 subunit together with either wild-type or mutant GlyR β subunit. Patch-clamp electrophysiology was used to analyse the effects of the L285R, W310C and M177R mutations on GlyR function. Dose-response relationships confirmed that the agonist sensitivity for GlyRs containing either L285R or W310C mutant did not change significantly relative to wild-type α1β heteromers (Fig. 5.9a,b). However, the M177R mutant resulted in a rightward shift in the EC_{50} value for glycine (Fig. 5.9c) from 39 ± 4 μM for wild-type α1β GlyRs to 162 ± 3 μM for $\alpha1\beta^{M177R}$ GlyRs (Table 5.2). Spontaneous activity of $\alpha1\beta^{L285R}$-containing GlyRs could be inhibited by the pore-blocker picrotoxin (Fig. 5.9d). However, due to the resistance of heteromeric α1β GlyRs to picrotoxin (Pribilla *et al* 1992), this inhibition was typically only about 100 pA. Single-channel recordings in the absence of glycine confirmed the spontaneous activity of $\alpha1\beta^{L285R}$ GlyRs (Fig. 5.9e). For $\alpha1\beta^{W310C}$ GlyRs, the maximal current amplitude was significantly reduced relative to wild-type GlyRs (Fig. 5.9f), suggesting that the number of heteromeric channels expressed in the plasma membrane was reduced.

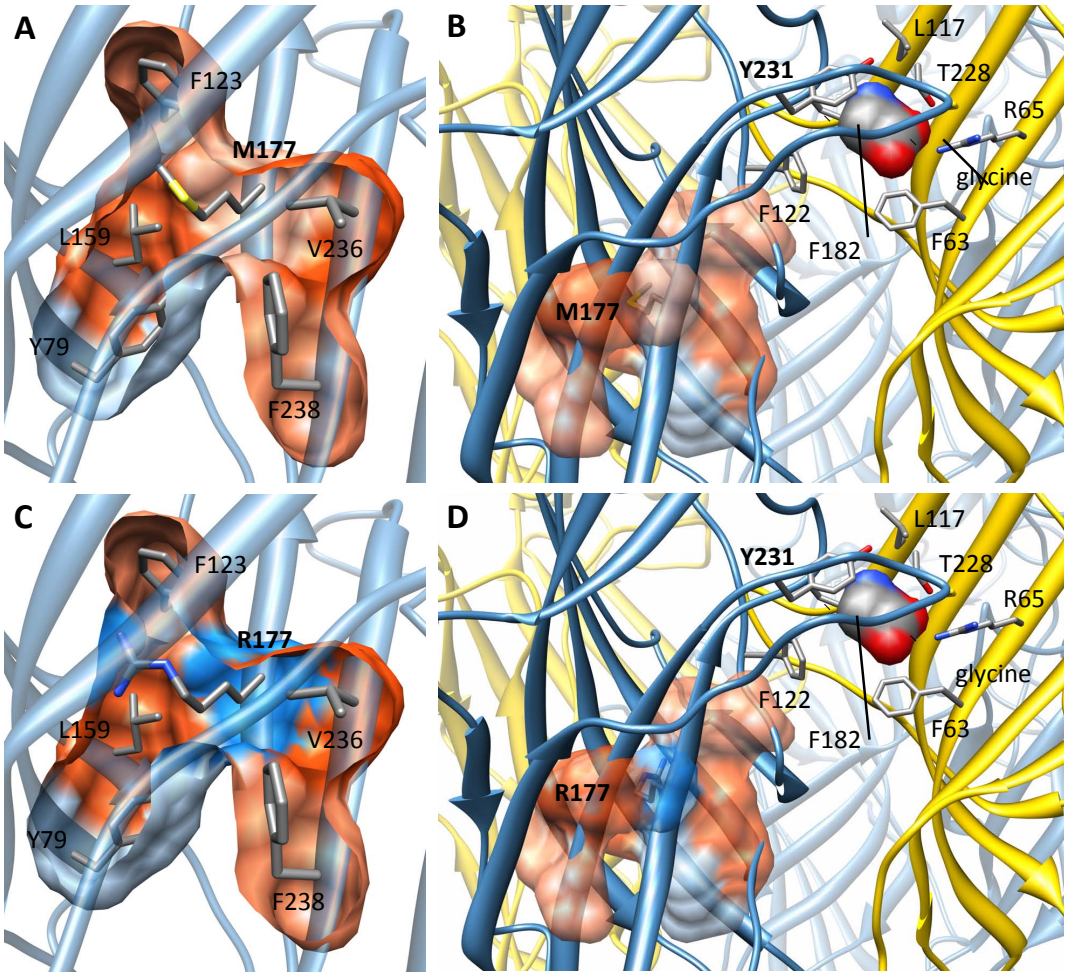

Figure 5.8. M177R disrupts local structure in the N-terminal extracellular domain. A: Side view of part of the GlyR β subunit extracellular domain showing the location of M177 within a β-sheet structure. The polarity of the surface residues is shown with hydrophobic residues depicted in orange, hydrophilic in blue and white indicating neutral polarity. **B**: M177 is located close to F182, which is involved in a cation-π interaction with the amide group of glycine. **C**: Mutation M177R is predicted to result in the loss of a hydrophobic side-chain and replacement with a hydrophilic residue with a positively-charged side chain. **D**: The M177R substitution is predicted disturb the local fold of the extracellular domain and have a knock-on effect for positioning of the critical ligand-binding residue F182.

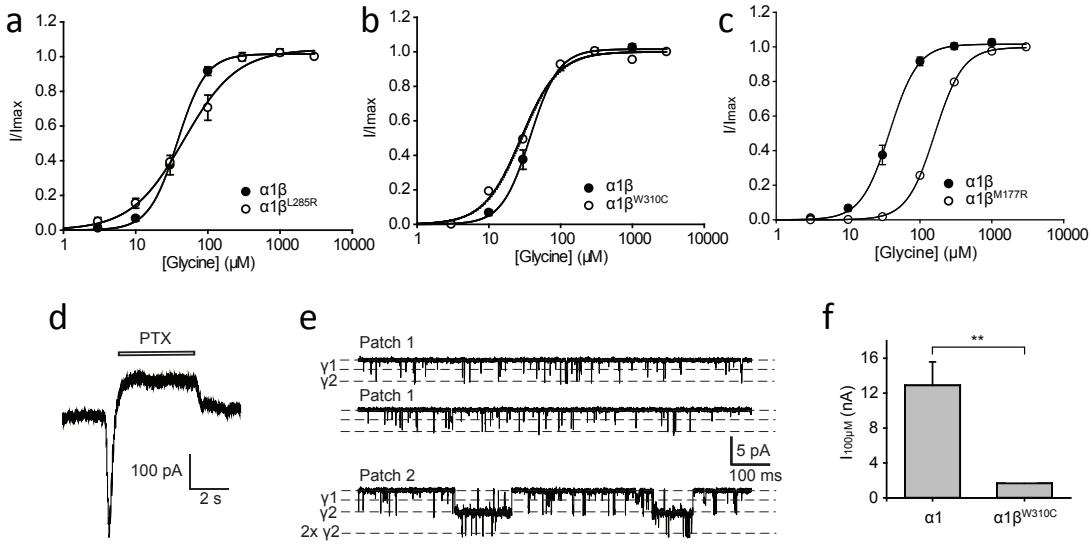

Figure 5.9. Functional characterization of GlyRs containing β^{L285R}, β^{W310C} and β^{M177R} using patch-clamp electrophysiology. **A - C**: Sample glycine dose-response traces in α1β, $\alpha1\beta^{L285R}$, $\alpha1\beta^{W310C}$ and $\alpha1\beta^{M177R}$ GlyRs. **D**: Inhibition of leak currents in hGlyR $\alpha1\beta^{L285R}$ by 100 μM picrotoxin (PTX). The inward current transient was always observed in 5 cells. **E:** Spontaneous single-channel currents for $\alpha1\beta^{L285R}$ GlyRs. Most openings were of 2 - 10 ms duration as illustrated in Patch 1. Occasionally, with multiple channels in the patch, longer activations were also observed, as in Patch 2. **F**: Mean currents activated by 100 μM glycine in cells expressing $\alpha_1\beta$ GlyRs in a 1:1 ratio relative to those in cells expressing $\alpha_1\beta^{W310C}$ GlyRs in a 1:10 ratio. p-values were calculated relative to GlyR $\alpha_1\beta$ heteromers by unpaired t-test: ** $p < 0.01$.

	Whole-cell patch-clamp electrophysiology			
	EC_{50} (μM)	**n_H**	**I_{max} (nA)**	**n**
α1β	39 ± 4	2.4 ± 0.2	14 ± 3	5
$\alpha1\beta^{L285R}$	49 ± 7	1.2 ± 0.1**	9 ± 4	3
$\alpha1\beta^{W310C}$	28, 58	1.7, 4.0	0.22, 0.15	2
$\alpha1\beta^{M177R}$	162 ± 3***	2.2 ± 0.1	8 ± 1	5

Table 5.2. Properties of α1β, $\alpha1\beta^{L285R}$, $\alpha1\beta^{W310C}$ and $\alpha1\beta^{M177R}$ GlyRs using whole-cell patch-clamp electrophysiology. EC_{50} values, Hill coefficients (n_H) and the maximal currents (I_{max}) are represented. p-values were calculated relative to the wild-type α1β subunit GlyRs with an unpaired t-test: ** $p < 0.01$. *** $p < 0.001$.

5.7. Discussion

In this chapter, I describe the identification and functional characterization of novel mutations in the GlyR β subunit gene (*GLRB*) that have revealed new pathogenic mechanisms underlying startle disease/hyperekplexia. We identified two new missense variants in *GLRB* causing L285R and W310C substitutions in membrane spanning domains TM2 and TM3, respectively. Using molecular modelling, I was able to establish the likely pathogenic mechanisms for L285R and W310C, as well as a third mutation - M177R - recently reported in a large Saudi Arabian family with hyperekplexia. L285R is a *de novo* mutation, i.e. not found in either parent, which disrupts the integrity of the 9' leucine hydrophobic girdle. $\alpha1\beta^{L285R}$ GlyRs were associated with leak conductance that could be blocked by picrotoxinin, indicating tonic channel opening. Since a leak conductance associated with a dominant GlyR α1 subunit mutation (Y128C) was recently described by Chung *et al* (2010), it has spurred curiosity that this is an emerging theme in startle disease. Introduction of pore-lining 9' arginine residues might have been expected to affect single-channel conductance. However, we found that the dominant maximum conductance level of $\alpha1\beta^{L285R}$ GlyRs (49 pS) was similar to that of α1β GlyRs (44 pS).

On first impressions, W310C appeared to be a classical recessive hyperekplexia mutation, appearing to interfere with GlyR trafficking rather than affecting glycine sensitivity. However, detailed analysis revealed both novel pathogenic mechanisms and mode of inheritance for this missense mutation. Molecular modelling revealed that W310 is a key residue involved in a hydrophobic stack formed by aromatic residues in TM1, TM3 and TM4 that determines GlyR subunit stoichiometry. The W310C substitution in TM3 is predicted to act by destabilizing intramembrane packing of α-helices, a result that is in accord with our functional studies, which suggested compromised cell-surface expression of $\alpha1\beta^{W310C}$. However, we cannot rule out the possibility that reductions in channel open probability or single-channel conductance may have contributed to the reduced current-carrying capacity of these mutant receptors. The inheritance of this mutation is of significant clinical interest, since mild

startle symptoms were reported in both parents. This mutation is likely to represent a case of incomplete dominance, i.e. a mutation that has an intermediate effect in heterozygous carriers.

The remaining *GLRB* substitution we studied in detail was M177R, which both molecular modelling and functional studies suggest interferes indirectly with agonist binding, by disrupting a local β-sheet fold in the large extracellular domain. M177 is located just five residues away from F182, which participates in a cation-π interaction with glycine. Thus, the pathogenic mechanism for M177R resembles that of a previously reported mutation G229D, which also displayed a reduced EC_{50} for glycine (Rees *et al* 2002).

Based on our recent experience with *GLRA1* and *SLC6A5*, I consider it likely that a number of recessive SNVs remain to be discovered in *GLRB*. In this regard, I found that eleven additional potentially pathogenic variants in *GLRB* have recently been deposited in dbSNP and 1000 genomes project databases. Many of these variants are associated with next-generation sequencing studies such as the NHLBI Exome Sequencing Project (Tennessen *et al* 2012), (Appendix tables 2 and 3). It is notable that Sanger DNA sequencing may not have validated these variants, and the pathogenic relevance to startle disease remains unproven. However, based on the discovery of four new *GLRB* variants associated with hyperekplexia in this year alone, I recommend that *GLRB* should have equal status alongside *GLRA1* and *SLC6A5* in the molecular genetic diagnosis of startle disease.

6. DEFECTIVE GLYCINERGIC NEUROTRANSMISSION IN ZEBRAFISH

6.1. Background

In 1996, Granato and co-workers performed a genetic screen detecting behavioural phenotypes for motor functions in mutant zebrafish. 166 mutants were identified from this screen exhibiting embryonic motility (Granato *et al* 1996). They were categorised into 14 phenotypes, according to their responses to touch: no response, normal but reduced response, vigorous but abnormal response, or simultaneous, bilateral contractions. 103 of these mutants displayed abnormal birefringence of muscular fibres suggesting impairments in the nervous system, neuromuscular junction or functional components of muscle such as excitation-contraction coupling. A subset of these mutants have proven to be useful models for defective glycinergic neurotransmission. For example, the zebrafish *shocked* (*sho*) mutant is characterised by an abnormal response to tactile stimulation. Typically, a wild-type fish would respond to tactile stimulation with an 'escape response' consisting of a C-bend, a counter-turn, and a bout of rapid swimming. However, the *sho* mutant responds with a single large contraction of the trunk (Fig. 6.1) - the so-called 'twitch once' phenotype (Granato *et al* 1996; Cui *et al* 2005).

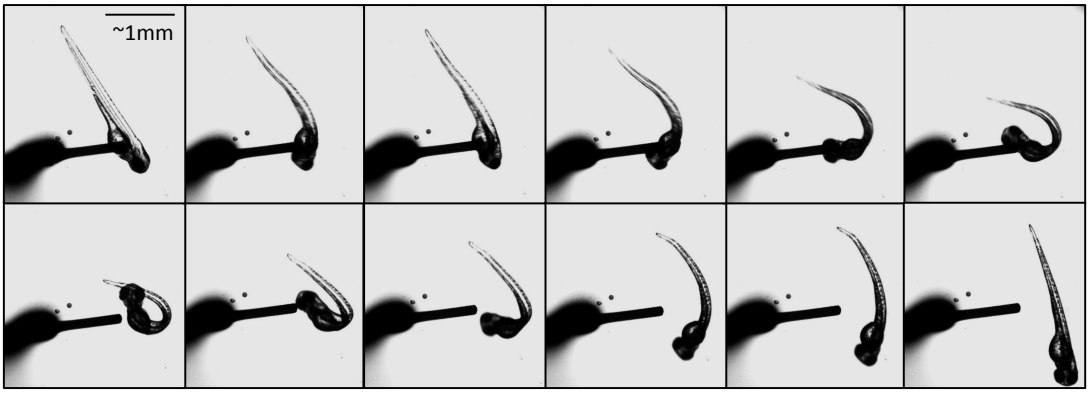

Figure 6.1. Movie stills of a *sho* mutant zebrafish responding to a tactile stimulus. A stimulus was applied to the yolk sack of a 72 hpf zebrafish *sho* mutant embryo. The *sho* mutant does not exhibit the classic escape behaviour seen in wild-type fish. Instead, a single large contraction of the trunk is seen. Single frames extracted from a high-speed movie provided by Julia Dallman.

As detailed in the introduction, this *sho* phenotype is caused by missense mutations in *slc6a9*, encoding GlyT1, although the mutation in the ta51e allele remains uncharacterised. The *sho* mutant has proven to be extremely useful in understanding homeostatic mechanisms that operate at glycinergic synapses – e.g. GlyRs were down-regulated in the *sho* mutant (Mongeon *et al* 2008), perhaps in an attempt to compensate for the excessive levels of glycine that would result when GlyT1 is inactivated. Despite these advances, a similar mutant affecting GlyT2 has remained elusive. However, a recent re-examination of zebrafish motility mutants uncovered a promising lead. The macho class mutant *schlaffi* (*sla*) maps between markers z1206 and z4999, a region that harbours *slc6a5*. However, whether this mutant harbours a mutation in *slc6a5* remained to be determined. Since no photographs or videos existed for these mutants, I was also unable to verify the accuracy of the *sla* phenotype described to date (see table 1.1).

Another useful class of zebrafish mutants was the so-called 'accordion' class (Table 6.1), which includes *accordion* (*acc*), *zieharmonika* (*zim*), *diwanka* (*diw*), *bandoneon* (*beo*), *quetschkommode* (*que*), *bajan* (*baj*) and *expander* (*exp*). Wild-type embryos at 24-28 hpf show characteristic coiling movements in response to tactile stimuli (Fig. 6.2). Applying the same stimuli to any of the 'accordian' mutant embryos induces simultaneous contractions of the trunk muscles on both sides of the embryo, resulting in a shortening of the body axis by 5-10%, followed by an extension back to the original size (Fig. 6.3) (Hirata *et al* 2004, 2010). This phenotype occurs as a result of the loss of an inhibitory mechanism that, in wild-type embryos, prevents motor neurons on one side from firing when the muscles on the other side are contracting. At first, it was demonstrated that most accordion mutants would represent defects in glycinergic transmission, since application of strychnine - a GlyR antagonist, to wild-type embryos phenocopied the abnormal 'accordion' behaviour (Granato *et al* 1996; Hirata *et al* 2004). However, it has since become apparent that there are a number of mechanisms that can result in an accordion-like phenotype and mutations have been discovered in genes involving cholinergic neurotransmission (e.g. acetylcholinesterase encoded by

ache, choline acetyltransferase, encoded by *chat*) or muscle function (e.g. the SERCA1 Ca^{2+} ATPase encoded by *atp2a1*). Other mutants, such as *expander* (*exp*) or *quetschkommod*e (*que*) remained uncharacterised for many years. However, the *que* mutation was characterised recently and shown to be a model of maple syrup urine disease (MSUD). This disorder is also known as branched-chain ketoaciduria, an autosomal recessive metabolic disorder affecting branched-chain amino acids. Curiously, the condition is named after the distinctive sweet odour of affected infants' urine. In humans, MSUD is caused by mutations in *BCKDHA*, *BCKDHB*, *DBT* and *DLD* which encode branched-chain keto acid dehydrogenase $E_1\alpha$, branched-chain keto acid dehydrogenase $E_1\beta$, dihydrolipoamide branched chain transacylase E2 and dihydrolipoamide dehydrogenase, respectively.

Together, these proteins make up the mitochondrial branched-chain α-keto acid dehydrogenase enzyme complex (BCKD) that is responsible for catalysing the branched-chain amino acids (BCAAs), isoleucine, leucine, and valine. Defective BCKD prevents normal breakdown, leading to a build up of these amino acids and their by-products in the body, which at high levels are toxic to the brain and other organs. This disease can be manageable by following a strict diet, but without control the accumulation of these substances ultimately leads to serious medical conditions such as brain damage and coma. In a recent study by Friedrich *et al* (2012) it was found that *que* mutant fish harbour a recessive mutation at the intron-exon boundary in *dbt*, resulting in a failure to remove the intron between exons 6 and 7. The intron contains multiple stop codons, which results in a prematurely terminated protein that is not functional. Elevated levels of BCAAs were also detected in the *que* zebrafish, confirming that a loss in *dbt* function results in abnormal BCAA metabolism and accumulation, similar to MSUD affected humans. *que* mutants represent a useful new model for MSUD and due to their relatively small nervous system with fewer cells, they provide a promising system to better characterize problems in the CNS as a result of BCAA toxicity (Friedrich *et al* 2012). This example illustrates quite nicely how characterisation of zebrafish ENU mutants can lead to useful models of human

neurological disease and provide models for the development of new pharmacotherapies.

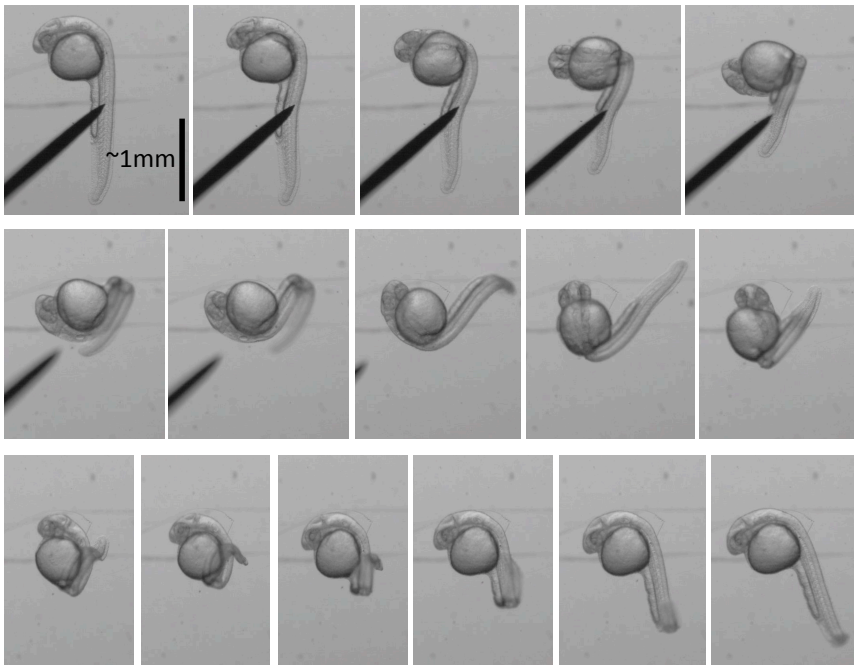

Figure 6.2. Wild-type zebrafish responding to tactile stimuli. A stimulus was applied to the body of a 22 hpf wild-type zebrafish embryo, which exhibits the classic coiling movements. Single frames were extracted from a high-speed movie provided by Hiromi Hirata.

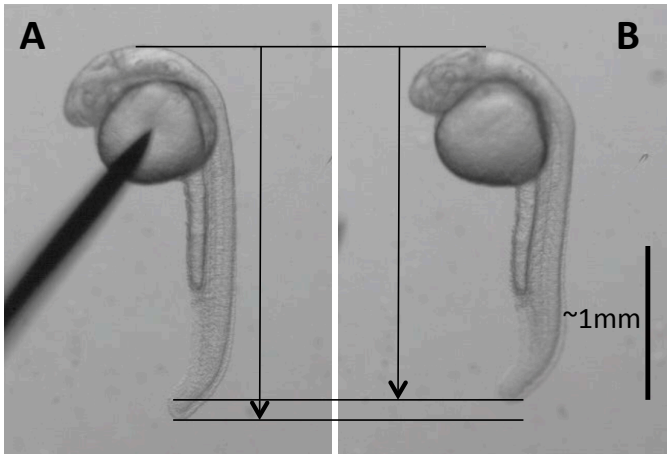

Figure 6.3. *beo* mutant (*accordion* class) zebrafish responding to tactile stimuli. A stimulus was applied to the yolk sack of a 22 hpf zebrafish *beo* mutant embryo (**A**). The *beo* mutant embryo does not exhibit coiling movements seen in wild-type fish. Instead, bilateral contractions of the trunk muscles are observed, leading to a shortening of the body axis (**B**). Single frames were extracted from a high-speed movie provided by Hiromi Hirata.

Mutant	Alleles	Mutation	Phenotype and gene defect	References
accordion (acc)	dta5 **mi25i** **mi289a** tc249a ti284a**†** tm286 tn218b tp72x **tq206§** ty20	Unknown **I97N** **T848I** Unknown Unknown Unknown Unknown Unknown **S766F** Unknown	Embryonic lethal. Touch-induced uncoordinated contraction of trunk muscles resulting in a contracted wavy notochord, 10-20% shorter than wild type. **Mutations I97N, S766F, T848I in the skeletal muscle sarcoplasmic reticulum Ca^{2+}-ATPase SERCA1 gene (*atp2a1*) on chromosome 3.**	Granato *et al* 1996; Odenthal *et al* 1996; Hirata *et al* 2004; Gleason *et al* 2004; Masino & Fetcho 2005
bajan (baj)	**tf247**	**IVS2-2A>C**	Embryonic lethal. Uncoordinated contraction of trunk muscles, eventually completely immotile. **Intron 2 splice acceptor site mutation in the choline acetyltransferase gene (*chat*) on chromosome 13.**	Granato *et al* 1996; Odenthal *et al* 1996; Wang *et al* 2008
bandoneon (beo)	tu230‡ **tp221** tm115 tf242 ta92† **tw38f§** ta86d **mi106a**	Allele lost Y79X unknown unknown unknown **L255R** unknown **R275H**	Embryonic lethal. Touch-induced uncoordinated contraction of trunk muscles resulting in a contracted wavy notochord, slightly bent up, 10-20% shorter than wild type. **Mutations Y79X, L255R, R275H in the GlyR beta subunit gene (*glrbb*) on chromosome 14.**	Granato *et al* 1996; Odenthal *et al* 1996; Hirata *et al* 2005; Masino & Fetcho 2005
diwanka (diw)	**ts286** **tv205a** **tz290**	**Q608X** **IVS4-1A>G** **W447X**	Embryonic lethal. Touch-induced uncoordinated contraction of trunk muscles resulting in a contracted wavy notochord, slightly bent up, 10 to 20% shorter than wild type, small eyes and enlarged hindbrain ventricle. **Intron 4 splice acceptor site and nonsense mutations W447X and Q608X in the procollagen lysine 2-oxoglutarate 5-dioxygenase 3 gene (*plod3*) on chromosome 23.**	Granato *et al* 1996; Odenthal *et al* 1996; Zeller & Granato 1999; Zeller *et al* 2002; Schneider & Granato 2006
expander (exp)	tu12	Unknown	Embryonic lethal. Uncoordinated contraction of trunk muscles resulting in a contracted wavy notochord, 10- 20% shorter than wild type. **Unknown gene on chromosome 11.**	Granato *et al* 1996; Odenthal *et al* 1996; Geisler *et al* 2007
quetschko-mmode (que)	ti274	IVS6+1G>A	Embryonic lethal. Uncoordinated contraction of trunk muscles resulting in a contracted wavy notochord, 10-20% shorter than wild type. **Splice site mutations dihydrolipoamide branched-chain transacylase E2 (*dbt*) on chr 22.**	Granato *et al* 1996; Odenthal *et al* 1996; Geisler *et al* 2007 Friedrich *et al* 2012
ziehharmo-nika (zim)	**sb55** **tf222a** **tm205§** tm206‡	**S226N** **G198R** **Y139X** Allele lost	Embryonic lethal. Uncoordinated contraction of trunk muscles resulting in a contracted wavy notochord, 10-20% shorter than wild type. Eventually becoming completely immotile. **Mutations Y139X, G198R and S226N in the acetylcholinesterase gene (*ache*) on chromosome 7.**	Granato *et al* 1996; Odenthal *et al* 1996; Behra *et al* 2002; Downes & Granato 2004

Table 6.1. Gene locations, nomenclature and mutations/phenotypes for zebrafish *accordion* mutants. ‡Mutant lost, †Viable allele, §Strongest allele.

6.1.1. *bandoneon (beo)* – defects in the GlyR βb subunit gene

Although there are clearly many routes to the 'accordion' phenotype (Fig 6.3), *bandoneon* mutants are defective in glycinergic synaptic transmission. Mutations in *glrbb* were identified as the basis for the *beo* mutant phenotype, leading to defective GlyR β subunit function. The underlying mutations for some, but not all *beo* alleles were identified (Hirata *et al* 2005; Masino & Fetcho 2005) (Table 6.2). These resulted in either missense changes (L255R in TM1, L275R in TM2) or a nonsense mutation (Y79X). Since GlyR β subunit mutations in humans cause hyperekplexia, the zebrafish *beo* mutant is a potentially useful animal model for this disorder (Hirata *et al* 2005, 2010).

beo	tu230‡ **tp221** tm115 tf242 ta92† **tw38f§** ta86d **mi106a**	Allele lost Y79X unknown unknown unknown **L255R** unknown **R275H**	Embryonic lethal. Touch-induced uncoordinated contraction of trunk muscles resulting in a contracted wavy notochord, slightly bent up, 10-20% shorter than wild type. **Mutations Y79X, L255R and R275H in GlyR βb subunit gene (*glrbb*) on chr 14.**	Granato *et al* 1996; Odenthal *et al* 1996; Hirata *et al* 2005; Masino & Fetcho 2005; Hirata *et al* 2010

Table 6.2. *beo* mutant alleles and known causative mutations. ‡ Mutant lost. §Strongest allele. †Viable allele.

Since the advent of the zebrafish genome sequencing project in 2001 (http://www.sanger.ac.uk/modelorgs/zebrafish.shtml), the identification of the chemically induced mutations responsible for abnormal phenotypes has been much improved. In addition, new techniques for gene disruption have been developed. For example, the knockdown of genes using specifically designed antisense morpholino oligonucleotides is now a standard technique in many laboratories. Morpholinos are injected into recently fertilised embryos and can either block the translation or splicing of a targeted mRNA, interfering with the gene expression *in vivo*. Transgenesis has also proved to be a successful tool for the analysis of gene function in zebrafish. DNA can be injected directly into embryos and the transgenic fish will exhibit cell- or tissue specific induced expression of the injected genes. For increased efficacy, virus- or transposon-mediated transgenesis methods have been developed, which have enhanced the application of controlled expression methods such the GAL4/UAS system (Hirata *et al* 2010).

6.2. Study aims

There is much that remains to be discovered and understood about the glycinergic transmission in zebrafish. Four *bandoneon* alleles remain uncharacterized, and it remained to be determined whether *schlaffi* represents an authentic GlyT2 mutant. Identification of the mutations in these zebrafish could uncover some interesting mutations in GlyR/GlyT2, and shed light on the structure and function of these proteins. I was also intrigued by a genetrap line recently reported for the *glra4a* gene, which indicated expression in neurones in the spinal cord and hindbrain. Since the corresponding gene (*GLRA4*) is a pseudogene in humans (Simon *et al* 2004), I decided that it would be interesting to knock down this gene in zebrafish, to learn more about the biological role of the α4 subunit. In addition to morpholino knockdown, I also wished to find out whether it was possible to interfere with receptor function by expressing an α4a subunit in zebrafish containing a dominant hyperekplexia mutation – equivalent to R271Q in the GlyR α1 subunit. The aims of this study were:

- To identify the causative mutations of the uncharacterised zebrafish *bandoneon* alleles by PCR and sequencing of *glrbb* from mutant and homozygote DNA.
- To clone a full-length zebrafish *glra4a* cDNA and introduce an artificial R278Q mutation by site-directed mutagenesis.
- To create a homology model of zebrafish GlyR α4aβb and use this model to predict how the R278Q mutation in zebrafish GlyR α4aβb disrupts the receptor function.
- To describe the phenotype of GlyR α4a knockdown mediated by morpholinos and the R278Q mutant.
- To describe the phenotype of the *sla* mutant, to compare this with morpholino knockdown of GlyT2, and determine whether *sla* corresponds to a GlyT2 defect.

Results

6.3. Identification of new *glrbb* mutations in *beo* alleles ta86d, ta92, tm115 and tf242

Genomic DNA was extracted from fin clips of *beo* carrier zebrafish alleles ta86d, ta92, tm115 and tf242 obtained from the zebrafish Stock Center in Tübingen, Germany. The nine coding exons of *glrbb* were amplified by standard PCR for each allele. Primer sequences for each exon can be found in the Appendix in Table 3. PCR products were cloned using the TOPO TA cloning kit and then sequenced by Sanger sequencing. Single nucleotide polymorphisms were identified by alignment with the *glrbb* exon consensus sequences using Sequencher 5.0. Nucleotides that differed from the reference sequence in one or more of the clones were noted. Nucleotide changes that resulted in missense or nonsense mutations that were specific to the line under examination could represent the causative mutation. Unique point mutations were found in all four *beo* alleles, with three resulting in nonsense mutations and one in a missense mutation. Mutations in *beo* alleles tm115, ta86d and ta92 are nonsense mutations resulting in premature termination of the GlyR βb polypeptide at residues Q87, Y79 and K343 respectively (Table 6.3; Figure 6.4).

Mutant *beo* allele	**cDNA**	**Protein**	**Details**
ta86d	T303A	Y79X	297 AGACTACTATGGACTA**T→A**CGTGTCAACATCTTCC 74 E T T M D ***** R V N I F
ta92	A1093T	K343X	1078 GAGGCGGAGAGAGCA**A→T**AGATCGCCACCAAGGAG 265 E A E R A ***** I A T K E
tm115	C325T	Q87X	311 AACATCTTCCTGAGA**C→T**AACGCTGGAATGACCCC 82 N I F L R ***** R W N D P
tf242	T301G	Y79D	286 GAGACTACTATGGAC**T→G**ATCGTGTCAACATCTTC 74 E T T M D **D** R V N I F

Table 6.3. Mutations identified in remaining *beo* alleles and the resulting changes in the GlyR βb subunit sequence.

These truncated proteins are predicted to be non-functional, since significant functional domains of the protein will be lost. Q87 and Y79 both lie in the large extracellular domain of the receptor. They both occur close to the N-terminus and therefore only a short part of the extracellular domain would be translated, with loss

of the entire TM1-TM4 region. K343X occurs within the large intracellular loop between TM3 and TM4. A truncation here would result in the loss of TM4 and a large majority of the intracellular loop, which is likely to have a dramatic effect on the tertiary fold of the subunit and a concomitant loss of the gephyrin binding site located between TM3 and TM4. The missense mutation identified in the tf242 allele results in Y79D substitution in the extracellular domain (Table 6.3; Fig. 6.4), which could affect ligand binding by neighbouring α subunits, or assembly of αβb heteromers. Taken together, these data indicate that mutations resulting in protein truncation are the predominant mechanism disrupting GlyR βb function in the *beo* allele series (4/7 alleles), followed by missense mutations affecting residues in the large extracellular domain (1/7 alleles) or TM1/TM2 (2/7 alleles).

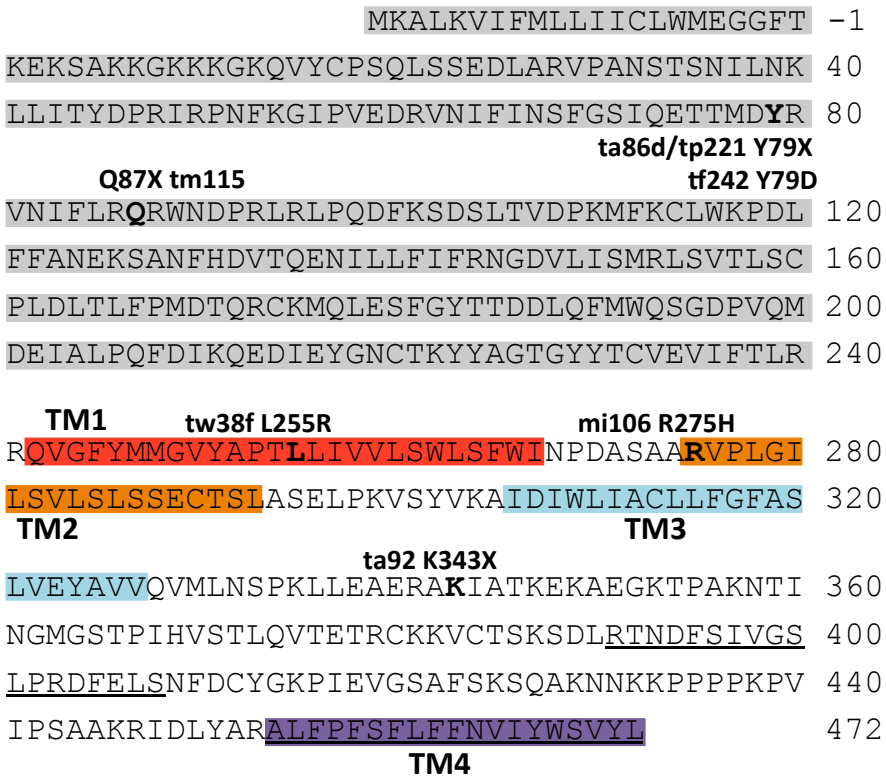

Figure 6.4. Zebrafish GlyR βb sequence indicating mutations found in different *beo* alleles. Sequence of the zebrafish GlyR βb (Uniprot I.D.: F1Q5M0) protein showing the residues affected by *beo* mutations. Predicted membrane-spanning domains (TM1-TM4) are indicated by coloured boxes, the extracellular domain is boxed in grey and the conserved gephyrin-binding motif is underlined.

6.4. A zebrafish genetrap reveals *glra4a* expression in brainstem and spinal cord neurones

A novel GAL4 gene-trap affecting the *glra4a* was discovered by Dr Max Suster at the Sars International Center for Marine Molecular Biology, Norway. In brief, *GAL4*, which encodes a yeast transcription factor activator protein, flanked by internal ribosome entry site (IRES) and a polyadenylation signal (Fig. 6.5) was integrated into *glra4a* between exon 1 and exon 2. GAL4 is a yeast protein and is not naturally present in other organisms, and has no effect, since it normally has no appropriate activating sequences to bind in the zebrafish. However, GAL4 works efficiently as a transcription factor when expressed in close proximity to an upstream activator sequence (UAS) that will usually be a fluorescent protein such as enhanced green fluorescent protein (EGFP). The zebrafish *glra4a*:GAL4 gene-trap line was therefore crossed with a UAS:EGFP line. The offspring express GAL4, which binds to the UAS and subsequently triggers the production of EGFP. Therefore, areas within the zebrafish where EGFP is detected indicate where *glra4a*:GAL4 is normally expressed.

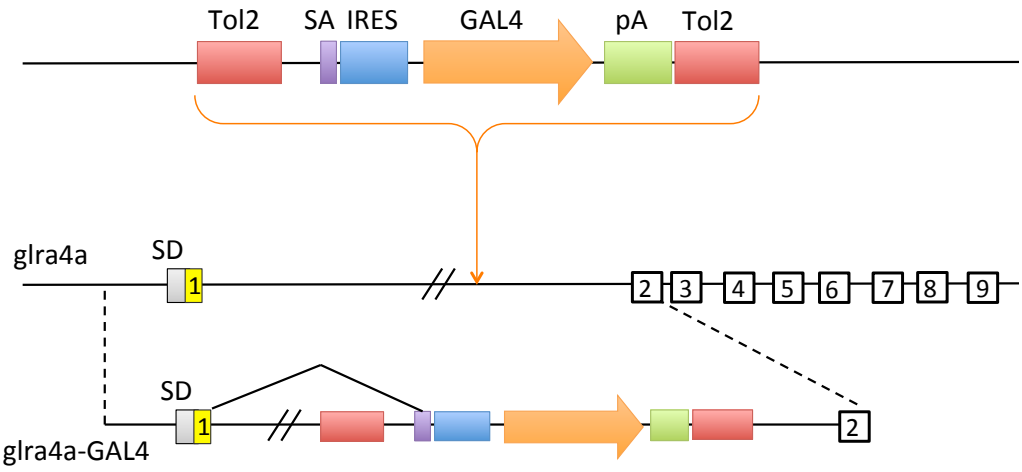

Figure 6.5. A schematic representation of creating a *glra4a* gene trap by gene transfer. The gene transfer vector Tol2 at either end of the DNA allows insertion into the zebrafish genome – in this case between exons 1 and exon 2 of *glra4a*. The splice acceptor (SA) interrupts normal splicing and ensures that the inserted sequence is transcribed. The internal ribosome entry site (IRES) is included to ensure that translation is active and the polyadenylation site (pA) interrupts the splicing and causes a stop in the translation, leading to a truncated protein.

EGFP expression directed by *glra4a*:GAL4 was detected in small interneuron populations in the hindbrain and spinal cord (Fig. 6.6). At 48 hpf expression of the gene was detected in commissural primary and secondary ascending neurons of the spinal cord, and at 72 hpf this expression was dramatically increased (Fig. 6.7).

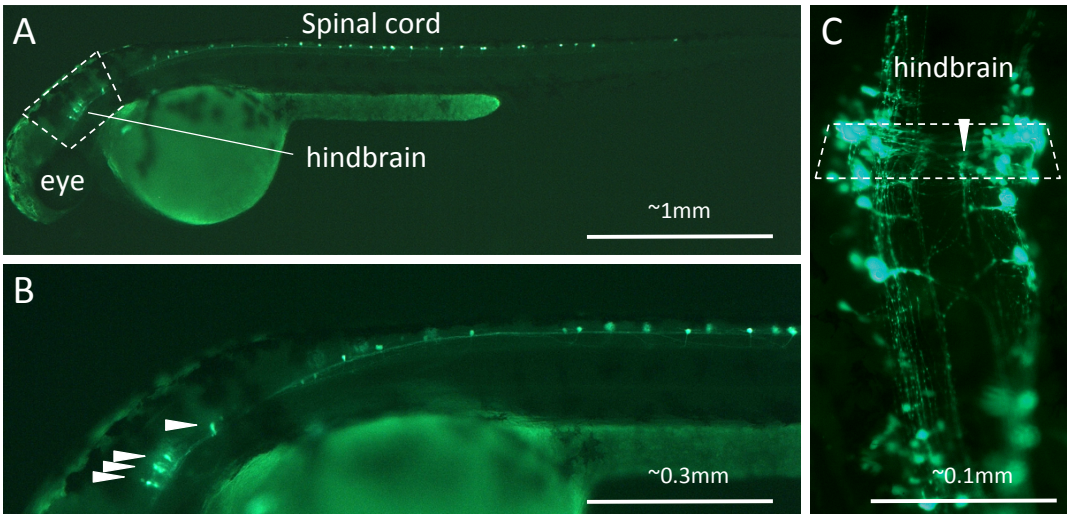

Figure 6.6. Localisation of GlyR α4a expression in the zebrafish brain and spinal cord. Expression of *glra4a:*GAL4 was determined by visualisation of EGFP. Expression is restricted to the spinal cord (**A**) and hindbrain (**B**), in commissural primary and secondary ascending neurons (**C** and Fig. 6.7). Images provided by Max Suster.

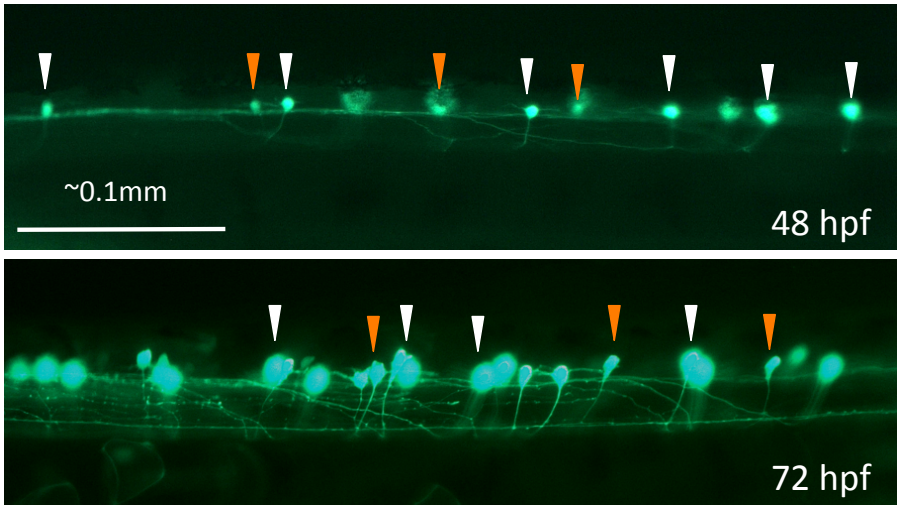

Figure 6.7. *glra4a*:GAL4 is expressed in commissural primary and secondary ascending neurons. The images show a portion of the zebrafish spinal cord at 48 hpf (upper panel) and 72 hpf (lower panel) expressing EGFP in the presence of *glra4a*:GAL4. CoPA neurons are indicated with a white arrow, and CoSA neurons with an orange arrow. Images provided by Max Suster.

6.5. Molecular cloning of the zebrafish GlyR α4a subunit cDNA and introduction of an artificial R278Q mutation by site-directed mutagenesis

The very first hyperekplexia-causing mutations to be discovered were two dominant missense mutations in *GLRA1* leading to substitutions in the same residue, R271Q and R271L (Shiang *et al* 1993). These mutations remain the most frequently reported dominant mutations and have been identified in a number of unrelated hyperekplexia patients (Schorderet *et al* 1994; Tijssen *et al* 1994; Elmslie *et al* 1996; Rees *et al* 2001; Kimura *et al* 2006; Chung *et al* 2010). Interestingly, mutations leading to R271X and R271P have also been reported (Gregory *et al* 2008; Lee *et al* 2012). In order to test the pathogenic mechanisms of the R271Q mutation, transgenic mice expressing the mutation were generated (Becker *et al* 2002. These mice display a dramatic phenotype with a clear exaggerated startle response, similar to that seen in the mouse *Glra1* and *Glrb* sporadic mutants *spasmodic* (Buckwalter *et al* 1994) and *spastic* (Kingsmore *et al* 1994; Mülhardt *et al* 1994), respectively (Becker *et al* 2002). R271 in the human GlyR α1 subunit is located in the extracellular loop that joins TM2 and TM3 in the subunit and disrupts the link between agonist binding and channel gating *in vitro* (Langosch *et al* 1994). Importantly, the R271Q mutant appears to be expressed normally at the cell surface (Chung *et al* 2010), where it can form heteromeric α1β subunit receptors.

I wished to discover whether introducing an equivalent mutation into a recombinant cDNA for a particular GlyR subunit and expressing the corresponding protein in zebrafish could yield a specific phenotype, comparing this to morpholino knockdown for the same gene. A full-length zebrafish *glra4a* (Uniprot I.D. Q5CZW5) cDNA in the pDNR-LIB vector (GeneService, Source Bioscience UK Ltd) was used as a template in PCR reactions with specifically designed primers to generate a *glra4a* cDNA with compatible restriction sites to allow their insertion into pCS2+ (*Glra4a*-EcoRI 5'-agagaattcgccaccatgctccctcaggtcataag-3' and *glra4a-Sal*I 5'-atggtcgacttaagggttggcgtggatatcctcgt-3'). Following PCR, the cDNA was digested with appropriate restriction enzymes and ligated into pCS2+. The constructs were then verified by Sanger DNA sequencing. I then used site-directed mutagenesis

(QuikChange, Agilent) to introduce the R278Q mutation involving two base changes (Fig. 6.8) into the zebrafish GlyR α4a subunit cDNA using specifically designed mutagenesis primers that contain the base changes highlighted in Fig. 6.8 (R278Q1 5'-ccacccagagctccggttcacaagcctcgctac-3' and R278Q2 5'-gtagcgaggcttgtgaaccggagctctgggtgg-3'). Once again, the constructs were verified by Sanger DNA sequencing to ensure that only the desired mutation had been introduced.

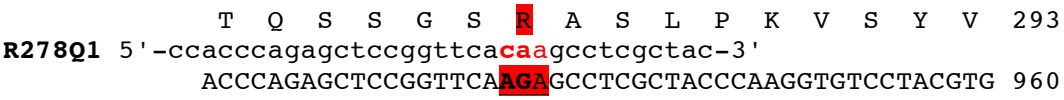

Figure 6.8. Mutagenesis of *glra4a* to introduce the R278Q missense mutation. Changes c.A941C and c.G942A (bold type) were introduced into zebrafish *glra4a* cDNA using site-directed mutagenesis, changing the codon and thus the transcribed residue from arginine to glutamine.

6.6. Morpholino knockdown and overexpression of an artificial GlyR α4a R278Q mutation reveals aberrant swimming behaviour

Splice site (SMO) and translation blocking (TMO) morpholinos α4a-SMO1, α4a-SMO2 and α4a-TMO (see Materials and Methods 2.1.12) were injected into zebrafish embryos. The morphology of the injected zebrafish was assessed at 48 hpf (Fig. 6.9) and did not reveal any gross anatomical changes apart from slightly smaller eyes that were apparent in embryos injected with α4a-TMO and α4a-SMO1 (Fig. 6.9, Left panel). This suggested that α4a-TMO and α4a-SMO1 were producing a consistent phenotype. Although it was not possible to monitor the efficacy of α4a-TMO knockdown, due to the lack of a specific GlyR α4a subunit antibody, RT-PCR was used to monitor the effects of α4a-SMO1 and α4a-SMO2. mRNA was extracted from zebrafish that had been injected with α4a-SMO1 and α4a-SMO2 plus non-injected zebrafish embryos (wild-type control). RT-PCR was performed using primers targeted within exon 6 (forward) and exon 8 (reverse) of zebrafish *glra4a* and the PCR products corresponding to exons 6-8 were run on an agarose gel (Fig. 6.9, upper right panel). This analysis suggested that binding of α4a-SMO1 to the splice acceptor site resulted in skipping of exon 7 in a proportion of transcripts (note lower 178 bp band), whilst binding of α4a-

SMO2 to the exon 7 acceptor site resulted in mis-splicing, so that a different 'acceptor' site was used within exon 7. Both outcomes result in a frameshift and loss of intact GlyR α4a subunits. RT-PCR with similar primers targeted against zebrafish *glra1* did not result in any aberrant splicing (Fig. 6.9, lower right panel) suggesting that the *glra4* morpholinos do not recognise this closely related target.

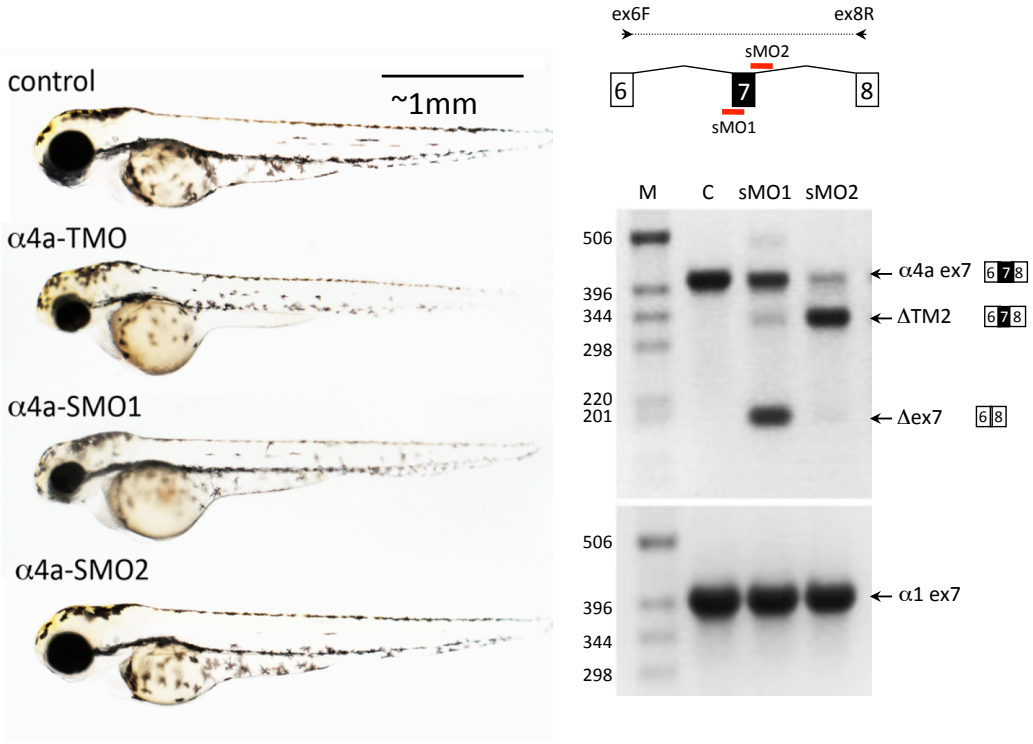

Figure 6.9. Morpholino oligonucleotides (SMO1 and SMO2) block splicing of *glra4a*. Left panel: gross morphology of wild-type embryos and those injected with GlyR α4a morpholinos at 72 hpf. Right panels: RT-PCR analysis of *glra4a* morpholinos confirmed the deletion of exon 7 for α4a-SMO1 (lane 3) and the deletion of part of exon 7 for α4a- SMO2 (lane 4). Intact GlyR α4a exon 7-9 PCR products are seen in the control (393 bp, lane 2) and smaller amounts are also observed in lanes 3 and 4. A fragment of around 178 bp is seen in sMO1, which is made up of exon 6 and 8 only. A fragment of 321 bp in SMO2 contains exons 6 and 8 and part of exon 7, which excludes the region that codes for TM2. GlyR α1 exon 7 is present in control and both SMO fish (lower panel). Photographic images provided by Max Suster.

Wild-type zebrafish embryos respond to tactile stimulus with an 'escape response' consisting of a C-bend, a counter-turn, and a bout of rapid swimming (Fig. 6.10, upper panels). However, behavioural analysis of zebrafish embryos injected with α4a-SMO1

gave rise to transient spasms and prolonged head retraction (Fig. 6.10, lower panels). Although the zebrafish appeared to eventually recover and swim away from the stimulus, this took significantly longer. Thus, knockdown with α4a-SMO1 significantly impairs escape behaviour. The same behaviour was observed in embryos injected with *in vitro*-transcribed RNA generated from my pCS2+-GlyR α4a subunit R278Q mutant (Fig 6.11). This suggests that morpholino knockdown by α4a-SMO1 and the GlyR α4a subunit R278Q mutant generate a similar phenotype.

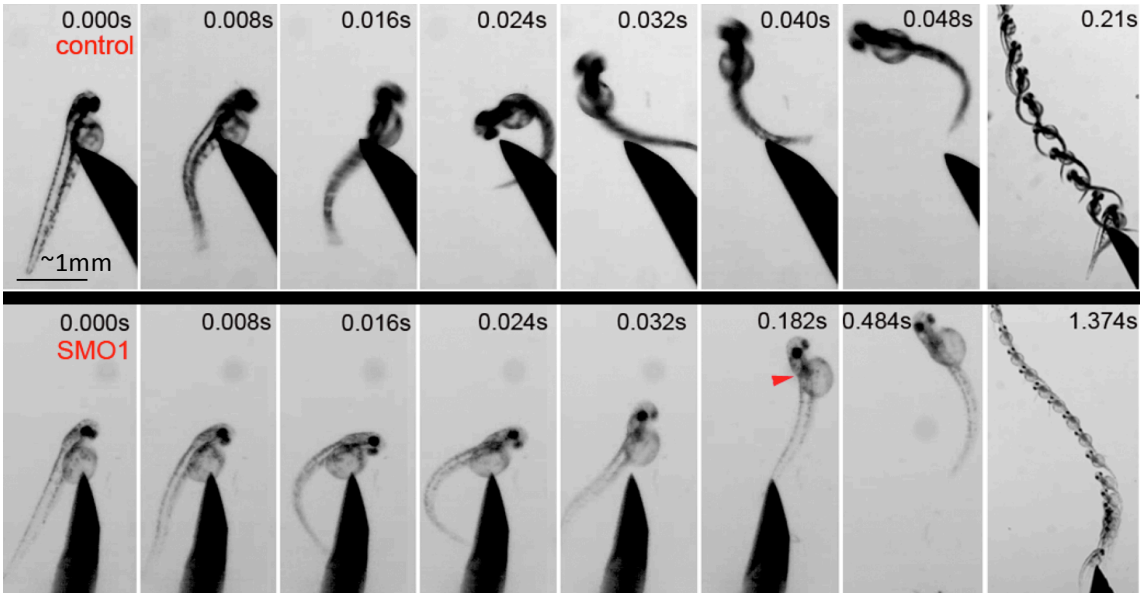

Figure 6.10. Embryos injected with α4a-SMO1 embryos fail to exhibit correct escape behaviour. The wild-type (control) fish responds to touch by with a C-bend, a turn away from the stimulus and rapid swimming. However, embryos injected with α4a-SMO1 respond more slowly, and have a pronounced head retraction - similar to a localised *beo*-like contraction (arrow). Single images were extracted from a high-speed movie of α4a-SMO1 morphants recorded by Max Suster.

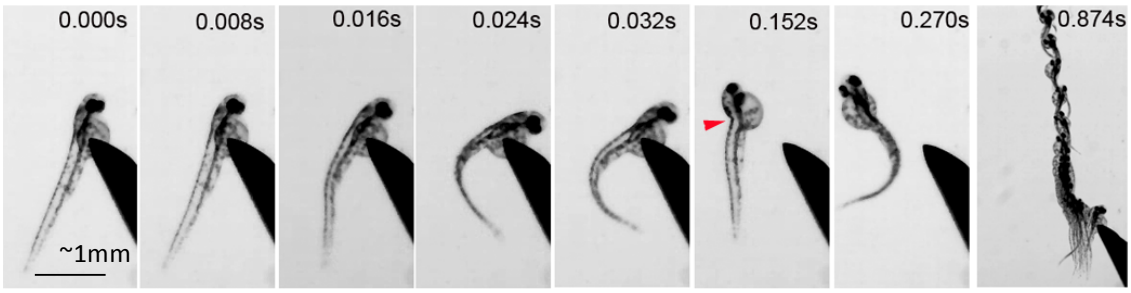

Figure 6.11. Embryos injected with mRNA for the α4a subunit R278Q mutant also fail to exhibit correct escape behaviour. Embryos injected with mRNA for the GlyR α4a subunit R278Q mutant escape more slowly, and have a pronounced head retraction - similar to a localised *beo*-like contraction (red arrow). Single images were extracted from a high-speed movie of α4a subunit R278Q morphants recorded by Max Suster.

6.7. Molecular modelling of the zebrafish α4aβb GlyR and GlyR α4a R278Q mutant

To gain further insights into the potential pathogenic mechanisms of the R278Q mutation, I constructed a homology model of the zebrafish α4aβb GlyR based on the crystal structure of the GluCl from *C. elegans* (PDB: 3RHW) (Hibbs & Gouaux 2011). A profile-profile alignment between the zebrafish GlyR α4a and βb subunits and the GluCl α subunit was generated using the MUSCLE and T-COFFEE web servers resulted in 37.7% (α4a) and 44.7% (βb) identity. A short part of the sequence was removed from both the α4a and the βb subunits at the N-terminus (residues 1-15 and 1-33, respectively), and the extended loop region between TM3 and TM4 of both α4a and βb subunits were removed (residues 319-396 and 335-450, respectively) since it was not possible to model these regions reliably due to very low sequence identity. The final alignments are shown in Fig. 6.12. The models were built with additional restraints enforcing disulphide bonds between cysteine residues involved in the cys-cys loops within each subunit (for GlyR α4a: residues C145-C159 and C205-C216, for GlyR βb: C160-C174 and C220-C232). Each model was assessed by MODELLER with the Discrete Optimized Protein Energy (DOPE) statistical potential score (Shen & Sali 2006) and the optimal model was selected based on the lowest score. The GlyR α4aβb homology model was calculated to have a normalized DOPE Z-score of -0.298, indicating a reasonable model, particularly considering the use of an alignment with less than 50% sequence identity.

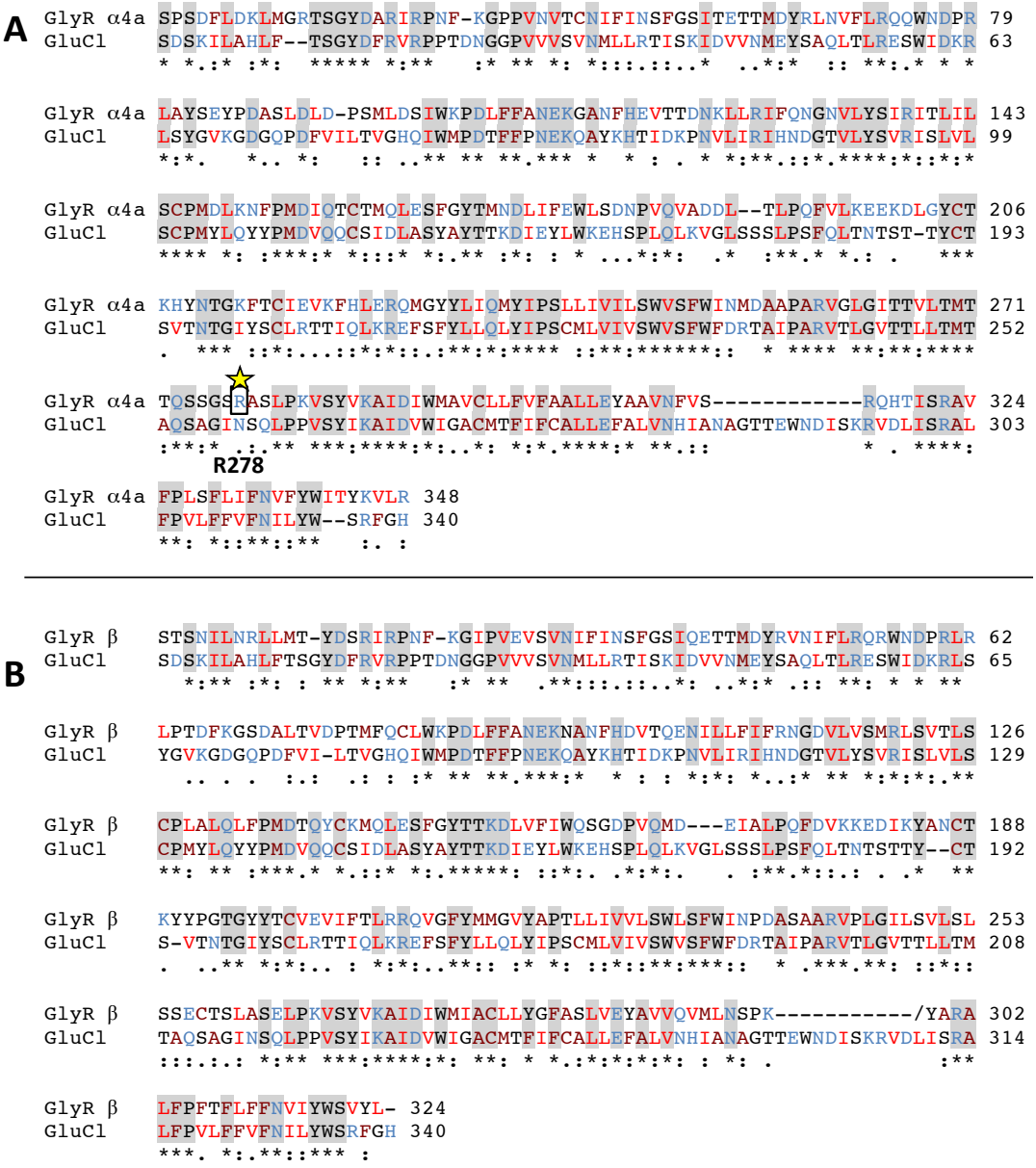

Figure 6.12. Alignments of zebrafish GlyR α4a (A) and βb (B) sequences with GluCl. Residues are coloured according to their hydrophobicity using the Kyte-Doolittle hydrophobicity scale (Kyte & Doolitte 1982), where the most hydrophobic residues are coloured in bright red (I, V, L), then less hydrophobic in dark red (F, C, A, M) and black (G, T, W, S, Y, P). Hydrophilic and charged residues are coloured in blue (D, E, N, Q, R, H, K). Residues highlighted in grey are conserved between GluCl and GlyR α4a (**A**) and βb (**B**) (also denoted by a *) and residues that are partially conserved (grouped by similar residue characteristics, such as charge, aromaticity, polarity and size) are indicated with a colon or stop. Gaps are indicated with a hyphen (-) and a forward slash (/) indicates a chain break.

In order to assess the accuracy of the model in more detail, the normalised DOPE score of each residue for a single GlyR α4a subunit and a single GlyR βb subunit from the best model were plotted alongside the corresponding residues in a single subunit of the GluCl template structure (Fig. 6.13). Peaks of high energy, particularly DOPE scores that are more than 0, are considered unfavourable. Areas in the plot where the energies overlap are usually correlated with regions of high conservation between target and template. This step was useful for highlighting specific regions of a model with high energies and therefore likely inaccuracies. Generally, there are no regions with DOPE scores above 1, and peaks of higher and lower DOPE scores were largely in agreement with those of GluCl, suggesting a native-like structure and a good model for both the GlyR α4a subunit (Fig. 6.13a) and the GlyR βb subunit (Fig. 6.13b).

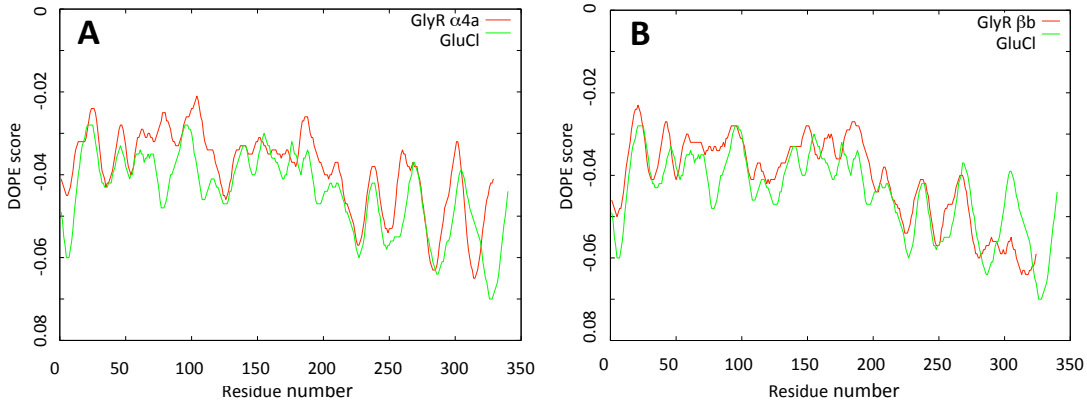

Figure 6.13. DOPE energy profiles of the zebrafish GlyR α4aβb model and the GluCl structure. The DOPE score for each residue of both the GlyR α4a subunit (**A**) and the GlyR βb subunit (**B**) from the GlyR α4aβb homology model, plotted against the residue number alongside DOPE scores for each residue of a single GluCl subunit.

Additional evaluation of the homology model of the entire GlyR α4aβb was performed using the QMEAN web server (http://swissmodel.expasy.org/qmean/cgi/index.cgi) (Benkert *et al* 2008, 2009). The QMEAN6 score of the model (Z-score = -2.8) fell within a range typically found for native proteins of similar size (Fig. 6.14) (Benkert *et al* 2009, 2011). Membrane proteins are expected to have low QMEAN Z-scores, since the scoring function is based on a sample of proteins taken from PDB, of which most are globular proteins. Membrane proteins are significantly more difficult to crystalize and

therefore there are a much smaller number of membrane proteins in the PDB relative to soluble ones. There are around 83,000 proteins in the PDB, of which only around 600 are membrane proteins according to SCOP, the Structural Classification of Proteins database (http://scop.mrc-lmb.cam.ac.uk/scop/).

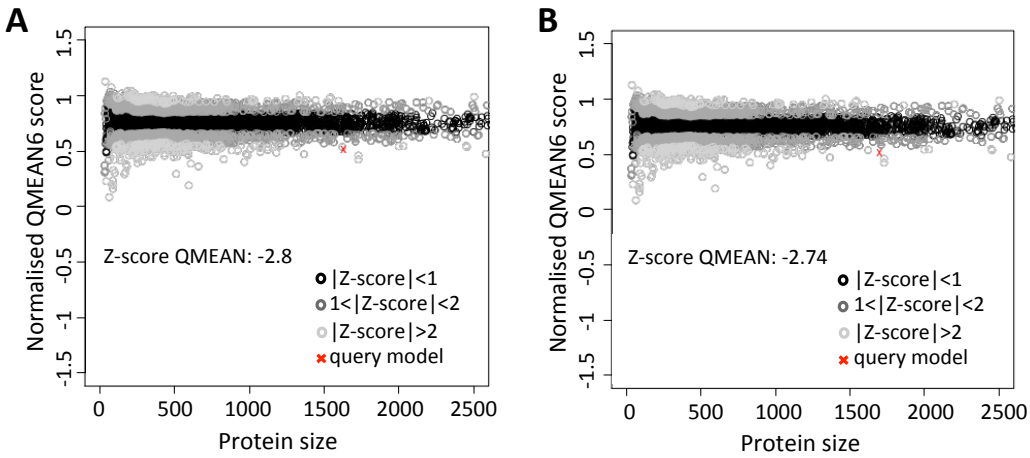

Figure 6.14. Normalised QMEAN6 scores for zebrafish GlyR α4aβb homology model (**A**) and the GluCl structure (**B**) plotted alongside other scores for a sample of crystal structures of other native proteins. The scores for both the GlyR α4aβb model and the GluCl structure lie on the border of the cluster of scores for native protein structures of a similar size. This indicates that the GlyR α4aβb homology model has a good normalised QMEAN score since it is close to the score of the native structure template, GluCl.

The R278Q substitution was modelled into the GlyR homology model with the '*swapaa*' command in Chimera (Pettersen *et al* 2004) using the Dunbrack backbone-dependent rotamer library (Dunbrack 2002) and taking into account the lowest clash score, highest number of H-bonds and highest rotamer probability (Fig. 6.15). R278 is predicted to be located on the second transmembrane domain helix (TM2), adjacent to the TM2-TM3 loop and close to the membrane in the extracellular space (Fig. 6.15a). The native, positively-charged arginine residue is substituted to a glutamine, which has an uncharged side-chain. A study by Lynch (2004) showed that the same startle disease-causing mutation in human GlyR α1 (R271Q) reduces the agonist glycine sensitivity.

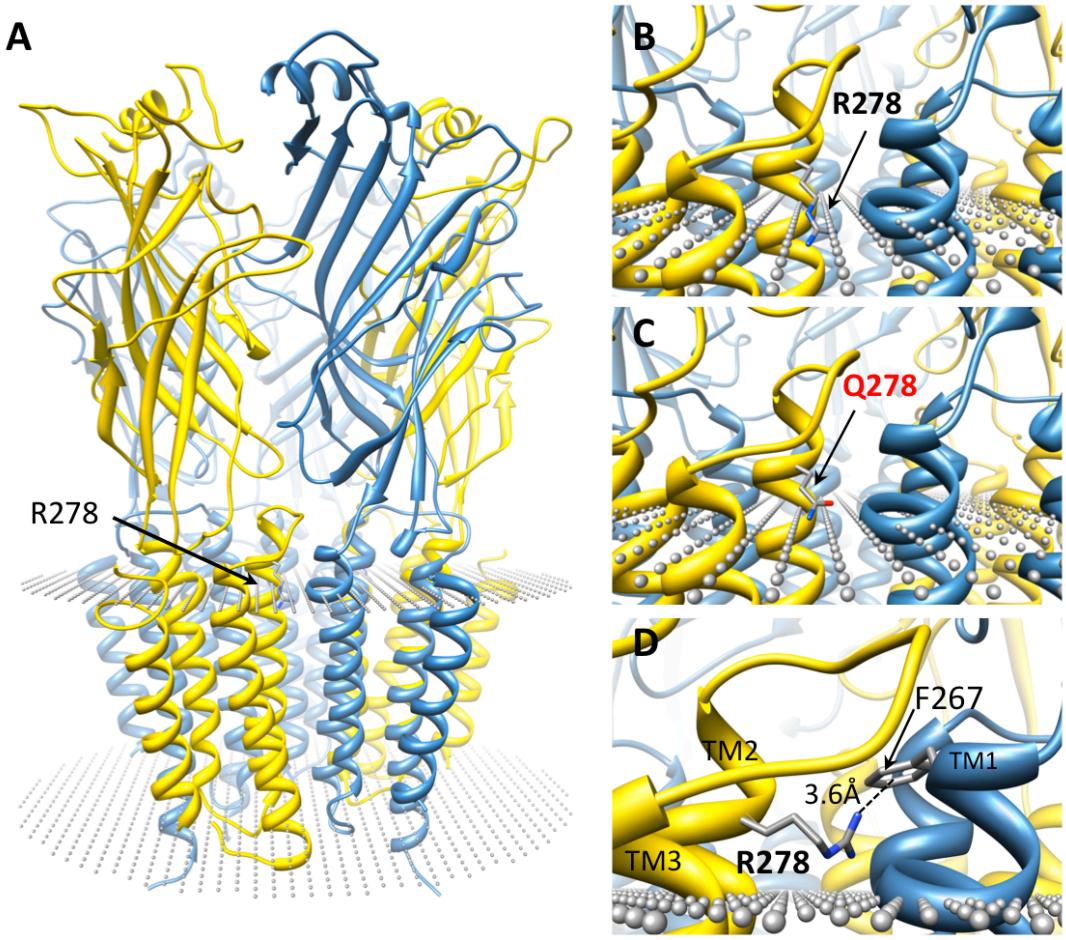

Figure 6.15. Molecular modelling of the zebrafish α4aβb GlyR and mutation R278Q. **A**: Side view of the zebrafish GlyR α4aβb, where α4a subunits are shown in yellow and βb subunits in blue. Grey spheres depict the predicted location of the membrane boundaries. **B**: R278 is predicted to be located on TM2 at the extracellular end of the helix, but close to the membrane. **C**: The residue is substituted to a glutamine that lacks the positive charge found in the arginine side chain. This may have some impact on the gating of the channel or interactions with adjacent subunits. **D**: An alternative rotamer of R278 shows a possible cation-π interaction with F267 located on the adjacent βb subunit. A mutation to Glu would abolish this potential interaction.

In addition, it has been suggested that this residue is located within the channel-gating pathway, which links agonist binding to the gating of the GlyR Cl^- channel (Pless *et al* 2007; Shan *et al* 2012). Although there are no obvious interactions or clashes that could provide a clear explanation for the structural effects of this mutation, it is possible that the change in charge at this position may have significant impact on the gating of the channel. Additionally, since the location of the residue is predicted to lie

close to a neighbouring subunit, it is also possible that a cation-π interaction is formed between R278 and F267 on TM1 of the adjacent GlyR βb subunit (Fig 6.15d). Therefore, substitution of this Arg to a Gln, will abolish this potential subunit-subunit interaction, which in turn may have an impact on the gating mechanism.

6.8. Identification of a new GlyT2 mutation in zebrafish in the *schlaffi* (*sla*) mutant

To ascertain whether a mutation in zebrafish *slc6a5* was responsible for the *schlaffi* (*sla*) phenotype, genomic DNA was extracted from homozygous *sla* mutants (allele th239) and siblings, kindly provided by Robert Giesler (Institute of Toxicology and Genetics, Karlsruhe Institute of Technology, Germany). Zebrafish *slc6a5* amplicons were generated from homozygous mutants by standard PCRs using the primers listed in the appendix (Table 4). SNVs were identified by aligning the sequences with native sequence contigs using Sequencher 5.0 software. SNVs found within each exon are listed in table 6.4.

slc6a5	**Amplicon size**	**SNV**	**Resulting codon**
Exon 1	151 bp	None	-
Exon 2	537 bp	None	-
Exon 3	264 bp	None	-
Exon 4	218 bp	G753A	A253A (synonymous)
Exon 5	283 bp	T819A	Y273X (nonsense)
Exon 6	232bp	None	-
Exon 7	226bp	None	-
Exon 8	245bp	None	-
Exon 9	192bp	C1377T G1419A	A459A (synonymous) L473L (synonymous)
Exon 10	198bp	None	-
Exon 11	150bp	C1985T	D565D (synonymous)
Exon 12	230bp	C1755T	Y585Y (synonymous)
Exon 13	195bp	C1896T	I632I (synonymous)
Exon 14	199bp	C1974G	G658G (synonymous)
Exon 15	248bp	G2097A	P699P (synonymous)
Exon 16	264bp	None	-

Table 6.4. SNVs found in *slc6a5* in *sla* th239 fish.

Eight synonymous SNVs, resulting in A253A, A459A, L473L, D565D, Y585Y, I632I, G658G and P699P substitutions and one SNV resulting in a nonsense mutation (Y273X) were identified in the sequencing of all 16 exons of *slc6a5* in *sla* homozygotes.

All SNPs result in synonymous changes (i.e. do not change the encoded amino acid) with the exception of T819A which represents a nonsense mutation in exon 5, resulting in a protein truncation (Y273X) within TM3. This mutation was detected in the heterozygous state in some siblings, but not in others, suggesting that it could be the causative mutation for the *sla* phenotype.

6.9. Morpholino knockdown and *sla* mutants reveal a *bandoneon*-like phenotype

For further proof that the mutation Y273X could underlie the *sla* phenotype, we designed a translation-blocking morpholino against the zebrafish GlyT2 mRNA and revived the *sla* mutant with the help of the Zebrafish Stock Center in Tübingen, Germany. In collaboration with Hiromi Hirata (Center for Frontier Research, National Institute of Genetics, Japan) I demonstrated that the *sla* mutant and microinjection of the GlyT2-TMO resulted in an 'accordion' phenotype, similar to the *beo* mutant. At 72 hpf, instead of exhibiting a classical escape response, they respond by simultaneous contraction of the muscles of the trunk leading to a shortening of the body, without swimming (Fig. 6.13). This phenotype mimics that of the *beo* mutants that harbour mutations in *glrbb*, encoding the GlyR βb subunit, but differs from the phenotype of GlyR α4a morpholino knockdown or R278Q mutant fish.

6.10. Discussion

Resequencing of the remaining *beo* mutant alleles resulted in the identification of missense and nonsense mutations in *glrbb*. Mutations Y79X, Y79D, Q87X and K343X were identified in alleles ta86d, tf242, tm115 and ta92, respectively (Table 6.5). Characterisation of these variants and their functional consequences is of significant biological interest, since the different *beo* alleles have different apparent strengths. For example, *beo* larvae typically die 7 dpf, either due to the inability to swim and feed effectively, or due to cumulative notochord damage from muscle contractions. It is interesting to note that a K343X nonsense mutation that was found in *beo*^{ta92} is apparently viable (Table 6.5). It is also noteworthy that both tp221 and ta86d alleles harbour the same mutation - Y79X (Table 6.5).

beo	tu230‡ **tp221** **tm115** **tf242** **ta92†** **tw38f§** **ta86d** **mi106**	Allele lost **Y79X Q87X** **Y79D** **K343X** **L255R** **Y79X** **R275H**	Embryonic lethal. Touch-induced uncoordinated contraction of trunk muscles resulting in a contracted wavy notochord, slightly bent up, 10-20% shorter than wild type. **Mutations Y79X, Q87X, Y79D, L255R, R275H in GlyR β subunit gene (*glrbb*) on chromosome 14.**	Granato *et al* 1996; Odenthal *et al* 1996; Hirata *et al* 2005; Masino & Fetcho 2005

Table 6.5. Updated table listing all mutations identified in *beo* alleles and the resulting changes in the amino acid sequence. ‡ Mutant lost. §Strongest allele. †Viable allele.

This unique data set will allow us to study the consequences of *beo* mutations on the functional properties of recombinant zebrafish GlyR α1βb GlyRs expressed in HEK293 cells or *Xenopus laevis* oocytes using patch-clamp methodologies. Whilst the function consequences of some *beo* mutations are evident (e.g. early protein truncations such as Y79X) future studies should focus on the missense changes Y79D, L255R and R275H to determine whether they affect GlyR α1βb trafficking, agonist binding, or the allosteric pathways linking agonist-binding to channel gating. Future work should also confirm that the ta92 allele is viable, since the K343X nonsense mutation results in a protein truncation and the loss of the majority of the intracellular loop EL2 (including the gephyrin binding motif) and TM4. Complementation of a truncated GlyR α1 subunit by an independently-expressed C-terminal domain containing TM4 was recently reported for in the mouse *oscillato*r mutant (Villmann *et al* 2009). It is therefore possible that the truncated GlyR βb protein is somehow incorporated into functional GlyRs, perhaps by domain swapping with other GlyR subunits, thereby mitigating lethality.

My studies into the zebrafish GlyR α4 subunit have revealed that this receptor subtype is expressed in subsets of hindbrain commissural and spinal neurons. The use of splice-site morpholinos revealed that knockdown of the GlyR α4 subunit gave rise to transient spasms and prolonged head retraction, resulting in aberrant escape behaviour. The same behaviour was observed in recombinant expression of a GlyR α4a

subunit R278Q mutant, mimicking the expression of a dominant hyperekplexia mutant. This observation suggests that incorporation of similar mutations into other zebrafish GlyR subunits could be used to generate a library of expression constructs that could be used to dissect the different biological functions of GlyR subtypes. It also suggests that the recombinant α4a subunit R278Q mutant was correctly processed, assembled into specific GlyR subtypes and transported to appropriate synapses. The head retraction observed could also be thought of as a partial, localised 'accordion'-like phenotype, suggesting that knockdown of multiple GlyR α subunits (e.g. α1, α2 and α4a) might be required to reproduce the *beo* phenotype.

The discovery of a nonsense mutation in exon 4 of *slc6a5* in the zebrafish *sla*^{th239} allele (resulting in a Y273X truncation) strongly suggests that a defect in the GlyT2 gene underlies the *sla* phenotype. However, the reported phenotype of the *sla* mutant did not match with the expected phenotype, i.e. that a GlyT2 mutant should resemble the *beo* phenotype. Since no photographs or videos of *sla* mutants exist, the *sla* mutant was revived and used to establish that the *sla* mutant and *slc6a5* knockdown did result in an 'accordion'-like phenotype (Fig. 6.16). Future work should seek to establish the causative mutation in the second *sla* allele, ty112, as well as attempting recovery of the *sla* phenotype by recombinant expression of zebrafish or human GlyT2. Certainly, the discovery of a new GlyT2 mutant will allow zebrafish researchers to investigate homeostatic changes that occur on loss of GlyT2 and perhaps as a screening tool to compare new and existing pharmacotherapies for hyperkplexia.

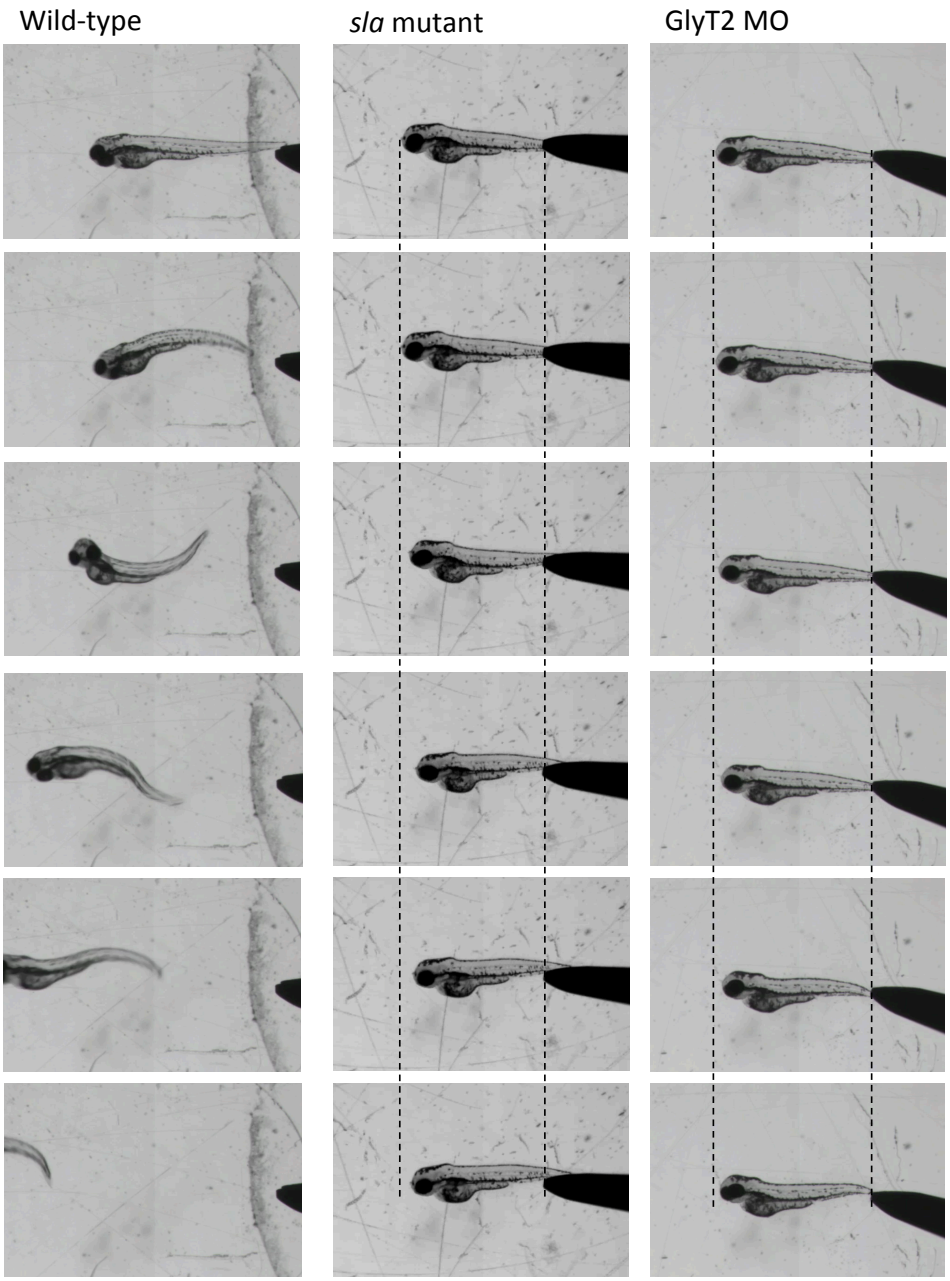

Figure 6.16. *sla* mutant fish and a GlyT2 TMOs reveal a *beo*-like phenotype. At 72 hpf, wild-type fish swim away from the tactile stimulus applied to the tail. In both *sla* mutants and GlyT2 TMO fish, the stimulus results in bilateral contractions within the trunk of the fish leading to shortening of the body. Dotted lines have been added as a guide to highlight shortening in *sla* mutants and GlyT2 MO fish. Single images were extracted from high-speed movies of *sla* mutants and GlyT2 morphants recorded by Hiromi Hirata.

7. GENERAL DISCUSSION

7.1. Mutations in the GlyT2 gene (*SLC6A5*) are a second major cause of startle disease

The majority of published studies on startle disease assume that the disorder is predominantly caused by dominant mutations affecting residues in the TM2-TM3 domains of the GlyR α1 subunit. However, it is now becoming increasingly apparent that recessive mutations in *GLRA1* are more common on a population basis (Chung *et al* 2010). Moreover, both recessive and dominant mutations in *SLC6A5*, encoding the presynaptic glycine transporter GlyT2 are increasing in prevalence (Rees *et al* 2006; Carta *et al* 2012; Giménez *et al* 2012). In part, this situation may be historical: *GLRA1* was the first startle disease gene to be discovered in 1993 and mutations in this gene explain the majority of cases of dominant hyperekplexia (Shiang *et al* 1993, 1995; Rees *et al* 1994; Schorderet *et al* 1994; Tijssen *et al* 1994; Elmslie *et al* 1996; Milani *et al* 1996; Seri *et al* 1997; Saul *et al* 1999; Miraglia Del Giudice *et al* 2003; Lapunzina *et al* 2003; Kimura *et al* 2006; Poon *et al* 2006; Becker *et al* 2008; Gregory *et al* 2008; Kang *et al* 2008; Chung *et al* 2010; Lee *et al* 2012). By contrast, mutations in *GLRB* and *SLC6A5* were first reported in 2002 and 2006, respectively (Rees *et al* 2002, 2006). This 'head start' for *GLRA1* is also reflected in the NCBI gene-testing registry (http://www.ncbi.nlm.nih.gov/gtr/). Of the clinical laboratories worldwide offering screening for startle disease, seven screen for *GLRA1* mutations whereas only three offer screening for *GLRB* and two for *SLC6A5* mutations. However, this testing does not reflect the true prevalence of disease alleles. In chapter 3, I described the discovery and analysis of twenty new pathogenic sequence variants for GlyT2 in 17 index cases, a finding that firmly establishes *SLC6A5* as a second major startle disease gene (Carta *et al* 2012). GlyT2 mutations were found in individuals from the UK, Australia, Canada, France, Italy, Jordan, the Netherlands, Portugal, Spain and the USA, also establishing that startle disease is a global phenomenon. In addition, a recurrent dominant GlyT2 mutation - Y705C - was also described in several individuals from the UK and Spain (Giménez *et al* 2012). My molecular modelling studies were particularly useful in understanding how these new GlyT2 mutations could cause startle disease, since the vast majority of the mutations were non-functional in [3H]glycine uptake assays.

Although many loss of function mutations resulted in protein truncation, missense mutations were predicted to: i) affect the conformation of TM2 (L237P, P243T); ii) indirectly affect glycine (A275T, E248K) or Na^+ (A275T, S513I) binding; iii) directly affect Cl^- binding (S513I); iv) affect the conformational mobility of EL4 (F547S) and v) affect cation-π or accessory protein interactions with the intracellular TM10-TM11 loop (Y656H, G657A).

7.2. Impact of the GlyT2 mutation causing CMD2 in Belgian Blue cattle

Selective breeding of cattle and other livestock has become a key consideration for the economics of the livestock farming industry. Efficient breeding programs are implemented to increase profitability by selection of animals with desirable traits with a view to produce offspring with increased feed conversion efficiency (a measure of the efficiency of converting feed mass into increased body mass), milk production and fertility. 'Elite sires' are utilised, often using artificial insemination methods, where one sire with desirable genetic traits will be used to produce many offspring. Whilst these breeding programs are generally very successful in selecting for these desirable traits, the inbreeding can lead to the emergence of recessive genetic defects, which can ultimately result in reduced effective population sizes leading to economic issues as well as concerns with animal welfare. In chapter 4, I described novel methods for detecting a mutation in *SLC6A5* in congenital muscular dystonia type 2 (CMD2) in Belgian Blue cattle. Since this is a recessive defect, it is possible to screen unaffected animals to determine whether they are carriers of the mutation. Animals that are found to carry the CMD2 could be excluded from future breeding, resulting in elimination of the mutation in future generations. However, since elite sires may have other valuable traits, another possibility is to use diagnostic tests to guide breeding strategy, avoiding 'at risk' matings of two carrier animals. In practice, organisations such as the British Blue Society may decide to insist on testing for a number of known genetic disorders, before registering sires for breeding.

7.3. Novel molecular mechanisms of GlyR β subunit mutations in startle disease

At the start of this project, mutations in *GLRB* were considered to be a very rare cause of startle disease and were only reported in a single family (Rees *et al* 2002). This situation was also reflected in mouse models of startle disease, with four models of GlyR α1 subunit dysfunction reported to date – *oscillator*, *spasmodic*, *cincinatti* and *nmf11* (Buckwalter *et al* 1994; Ryan *et al* 1994; Holland *et al* 2006; Traka *et al* 2006) versus one model of GlyR β dysfunction – *spastic* (Kingsmore *et al* 1994; Mülhardt *et al* 1994). However, we suspected that other *GLRB* mutations remained to be discovered due to methodological considerations. Early attempts at screening *GLRB* used single-strand conformation polymorphism (SSCP) or bi-directional di-deoxy fingerprinting (ddF) methods (Milani *et al* 1998; Rees *et al* 2002), which have variable sensitivity for mutation detection. By contrast, our study and two other recent reports (Al Owain *et al* 2012; Lee *et al* 2012) used Sanger DNA sequencing - currently the method of choice for mutation detection - although this is likely to be superseded in time by targeted next-generation sequencing panels (Lemke *et al* 2012).

It is also clear that there are key differences in the relative damage caused by equivalent missense changes in GlyR α1 versus GlyR β subunits. For example, dominant startle disease mutations in *GLRA1* predominantly affect key amino acid residues in the TM2 domain and TM2-TM3 linker region of the GlyR α1 subunit, where they act to uncouple ligand binding from channel gating. Shan and colleagues (2011) recently demonstrated that GlyR function is less sensitive to hyperekplexia-mimicking mutations introduced into the TM2-TM3 loop of the GlyR β subunit than that of the α1 subunit. This suggests that the GlyR α1 TM2-TM3 loop dominates the β subunit in gating heteromeric α1β GlyRs and in turn that it is perhaps unlikely that a set of equivalent dominant mutations in the TM2-TM3 linker will be found in *GLRB*. Rather, my molecular modelling studies suggest that *GLRB* mutations found to date appear to: i) cluster near key glycine binding residues (M177R, G229D) or ii) are found in membrane-spanning domains TM1-TM3, where they affect ion-channel function (L285R) or disrupt hydrophobic side-chain stacking (W310C). Other potential

pathogenic mechanisms for *GLRB* include protein truncation (via deletion, frameshift, splice site or nonsense mutations) or mutations in the gephyrin-binding site located between TM3 and TM4. At least *in vitro*, artificial missense mutations affecting single amino acids, such as F398A in the GlyR β subunit gephyrin-binding motif are capable of significantly impairing GlyR β-gephyrin interactions (Kim *et al* 2006), which would in turn abolish GlyR clustering at synapses.

7.4. Zebrafish glycinergic transmission - novel mutations and phenotypes

I also studied GlyRs and GlyTs in zebrafish, an organism that is amenable to developmental and genetic analysis using mutagenesis, gene-traps and rapid targeted gene 'knockdown' using antisense morpholinos. Curiously, zebrafish possess seven different GlyR genes (*glra1*, *glra2*, *glra3*, *glra4a*, *glra4b*, *glrba* and *glrbb*), probably due to chromosome duplications early in the evolution of teleosts (Hirata *et al* 2010). This phenomenon is well known in zebrafish, and has key advantages for the study of gene function, since individual paired genes typically differ in terms of expression patterns and functional roles. For example, the zebrafish mutant bandoneon (*beo*) harbours mutations in one of the paired GlyR β subunit genes (*glrbb*), resulting in bilateral muscle contractions due to loss of reciprocal glycinergic inhibition of motor circuits (Hirata *et al* 2005). However, the other GlyR β subunit gene (g*lrba*) is unable to compensate for the loss of *glrbb* function, suggesting that it forms part of a distinct GlyR with a different function (Hirata *et al* 2005). Since mutations in the genes for the postsynaptic GlyR α1 and β subunits and the presynaptic glycine transporter GlyT2 all result in startle disease in humans, it was curious that no mutations in *glra1* or *slc6a5* that mimic *beo* had been discovered in zebrafish. This suggested that the *beo* phenotype might result from the loss of multiple GlyR subtypes, or that GlyR βb participates in embryonic GlyRs that do not contain GlyR α1. Some supporting evidence for the former possibility has been provided by my studies on the zebrafish GlyR α4a subunit gene. My initial interest in this subunit was stimulated by the discovery of a *glra4a* gene-trap, which demonstrated the expression of *glra4a* in brainstem and spinal cord neurons. Morpholino knockdown of *glra4a* or

overexpression of a recombinant GlyR α4a subunit harbouring a dominant startle disease mutation (equivalent to R271Q) resulted in defective escape behaviours and a localised *beo*-like phenotype. Taken together, these data suggest that GlyR α4a plays an important role at glycinergic synapses in the brainstem and spinal cord, mediating a key behavioural response. This is the first example of a phenotype associated with the GlyR α4 subunit, but does not explain why this gene is a pseudogene in humans. Lastly, I was able to demonstrate that the mutant *schlaffi* (*sla*) represents a GlyT2 mutant, caused by a nonsense mutation in *slc6a5*. Curiously, the phenotype originally described for this mutant is likely to be incorrect, since analysis of the mutant phenotype and *slc6a5* morpholino knockdown revealed a *beo*-like phenotype. This suggests that *sla* should be re-assigned to the accordion class of zebrafish mutants.

7.5. Future prospects

Based on the results presented in this thesis, I recommend that *GLRA1*, *GLRB* and *SLC6A5* should have equal status in the molecular genetic diagnosis of startle disease. Mutations in *GLRB* and *SLC6A5* were associated with additional clinical symptoms, such as severe neonatal apnoea episodes (Rees *et al* 2006; Carta *et al* 2012), which have additional implications for patient care. Fortunately, in most cases, a positive response was seen to treatment with clonazepam (Rees *et al* 2006; Carta *et al* 2012), meaning that startle disease responds well to pharmacotherapy. Although sequential genetic analysis of hyperekplexia cases has revealed many additional mutations in *GLRA1*, *GLRB* and *SLC6A5*, there are families that remain mutation-negative for these major genes of effect. This suggests that other startle disease genes exist or that a number of different mutations remain to be characterized in regulatory elements at existing loci, e.g. upstream promoters. However, it is not currently straightforward to address this possibility. For example, the minimal *SLC6A5* promoter encompasses >90 kb upstream of exon 1 and regulatory motifs remain uncharacterized (Zeilhofer *et al* 2005). Another possibility is that mutations exist in genes unrelated to GlyT2s or GlyTs. There are certainly precedents for extensive genetic heterogeneity in other neurological disorders. For example, mutations in both inhibitory $GABA_A$ receptor (Baulac *et al*

2001; Harkin *et al* 2002) and voltage-gated Na^+ channel genes (Wallace *et al* 1998; Escayg *et al* 2000) are associated with generalized epilepsy with febrile seizures plus (GEFS+). In my view, the future challenges are to: i) understand the regulatory elements controlling transcription of *GLRA1*, *GLRB* and *SLC6A5* and ii) use alternative disease gene-discovery techniques, such as array comparative genomic hybridization (CGH) and exome/genome sequencing to search for pathogenic mutations in novel startle disease genes.

It would also be interesting to explore the structures of GlyT2 and GlyRs further using additional computational methods. GlyT2 is known to bind 3 Na^+ ions, but the position of only two can be predicted confidently from homology modelling. Determination of the binding site of this third sodium may shed light on the pathogenic mechanisms of some mutations in GlyT2 that lead to hyperekplexia that remain unsolved. Molecular dynamics studies may also be very useful for understanding the flexibility in certain regions of the glycine transporters and receptors that contribute to their correct function. By highlighting critical residues involved in the flexibility of these proteins may also aid the understanding of further startle-disease related mutations.

REFERENCES

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR (2010) A method and server for predicting damaging missense mutations. *Nat Methods* 7: 248-249.

Akiyama M (2006) Pathomechanisms of harlequin ichthyosis and ABCA transporters in human diseases. *Arch Dermatol* 142: 914-918.

Al-Futaisi AM, Al-Kindi MN, Al-Mawali AM, Koul RL, Al-Adawi S, Al-Yahyaee SA (2012) Novel mutation of *GLRA1* in Omani families with hyperekplexia and mild mental retardation. *Pediatr Neurol* 46: 89-93.

Al-Owain M, Colak D, Al-Bakheet A, Al-Hashmi N, Shuaib T, Al-Hemidan A, Aldhalaan H, Rahbeeni Z, Al-Sayed M, Al-Younes B, Ozand PT, Kaya N (2012) Novel mutation in *GLRB* in a large family with hereditary hyperekplexia. *Clin Genet* 81: 479-484.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215: 403-410.

Bajaj K, Madhusudhan MS, Adkar BV, Chakrabarti P, Ramakrishnan C, Sali A, Varadarajan R (2007) Stereochemical criteria for prediction of the effects of proline mutations on protein stability. *PLoS Comput Biol* 3: e241.

Bakker MJ, van Dijk JG, van den Maagdenberg AMJM, Tijssen MAJ (2006) Startle syndromes. *Lancet Neurol* 5: 513-524.

Baulac S, Huberfeld G, Gourfinkel-An I, Mitropoulou G, Beranger A, Prud'homme JF, Baulac M, Brice A, Bruzzone R, LeGuern E (2001) First genetic evidence of $GABA_A$ receptor dysfunction in epilepsy: a mutation in the γ2-subunit gene. *Nat Genet* 28: 46-48.

Becker K, Breitinger HG, Humeny A, Meinck HM, Dietz B, Aksu F, Becker CM (2008) The novel hyperekplexia allele *GLRA1*(S267N) affects the ethanol site of the glycine receptor. *Eur J Hum Genet* 16: 223-228.

Becker K, Hohoff C, Schmitt B, Christen H-J, Neubauer BA, Sandrieser T, Becker C-M (2006) Identification of the microdeletion breakpoint in a *GLRA1* null allele of Turkish hyperekplexia patients. *Hum Mutat* 27: 1061-1062.

Becker L, von Wegerer J, Schenkel J, Zeilhofer HU, Swandulla D, Weiher H (2002) Disease-specific human glycine receptor α1 subunit causes hyperekplexia phenotype and impaired glycine- and $GABA_A$-receptor transmission in transgenic mice. *J Neurosci* 22: 2505-2512.

Behra M, Cousin X, Bertrand C, Vonesch JL, Biellmann D, Chatonnet A, Strähle U (2002) Acetylcholinesterase is required for neuronal and muscular development in the zebrafish embryo. *Nat Neurosci* 5: 111-118.

Bellini G, Miceli F, Mangano S, Miraglia del Giudice E, Coppola G, Barbagallo A, Taglialatela M, Pascotto A (2007) Hyperekplexia caused by dominant-negative suppression of GlyR α1 function. *Neurology* 68: 1947-1949.

Ben-Yona A, Bendahan A, Kanner BI (2011) A glutamine residue conserved in the

neurotransmitter sodium symporters is essential for the interaction of chloride with the GABA transporter GAT-1. *J Biol Chem* 286: 2826-2833.

Benkert P, Biasini M, Schwede T (2011) Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* 27: 343-350.

Benkert P, Künzli M, Schwede T (2009) QMEAN server for protein model quality estimation. *Nucleic Acids Res* 37: 510-514.

Benkert P, Tosatto SCE, Schomburg D (2008) QMEAN: A comprehensive scoring function for model quality assessment. *Proteins* 71: 261-277.

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The Protein Data Bank *Nucleic Acids Res* 28: 235-242.

Bill BR, Petzold AM, Clark KJ, Schimmenti LA, Ekker SC (2009) A primer for morpholino use in zebrafish. *Zebrafish* 6: 69-77.

Bocquet N, Nury H, Baaden M, Le Poupon C, Changeux J-P, Delarue M, Corringer P-J (2009) X-ray structure of a pentameric ligand-gated ion channel in an apparently open conformation. *Nature* 457: 111-114.

Boehm S, Harvey RJ, Holst Von A, Rohrer H, Betz H (1997) Glycine receptors in cultured chick sympathetic neurons are excitatory and trigger neurotransmitter release. *J Physiol* 504: 683-694.

Bormann J, Rundström N, Betz H, Langosch D (1993) Residues within transmembrane segment M2 determine chloride conductance of glycine receptor homo- and hetero-oligomers. *EMBO J* 12: 3729-3737.

Brejc K, van Dijk WJ, Klaassen RV, Schuurmans M, van Der Oost J, Smit AB, Sixma TK (2001) Crystal structure of an ACh-binding protein reveals the ligand-binding domain of nicotinic receptors. *Nature* 411: 269-276.

Bröer S (2006) The SLC6 orphans are forming a family of amino acid transporters. *Neurochem Int* 48: 559-567.

Brooks BR, Brooks CL 3rd, Mackerell AD Jr, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caflisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, Karplus M (2009) CHARMM: The biomolecular simulation program. *J Comput Chem* 30: 1545-1614.

Brune W, Weber RG, Saul B, von Knebel Doeberitz M, Grond-Ginsbach C, Kellerman K, Meinck HM, Becker CM (1996) A *GLRA1* null mutation in recessive hyperekplexia challenges the functional role of glycine receptors. *Am J Hum Genet* 58: 989-997.

Buckwalter MS, Cook SA, Davisson MT, White WF, Camper SA (1994) A frameshift mutation in the mouse α1 glycine receptor gene (*Glra1*) results in progressive neurological symptoms and juvenile death. *Hum Mol Genet* 3: 2025-2030.

Carta E, Chung SK, James VM, Robinson A, Gill JL, Remy N, Vanbellinghen JF, Drew CJ, Cagdas S, Cameron D, Cowan FM, Del Toro M, Graham GE, Manzur AY, Masri A,

Rivera S, Scalais E, Shiang R, Sinclair K, Stuart CA, Tijssen MA, Wise G, Zuberi SM, Harvey K, Pearce BR, Topf M, Thomas RH, Supplisson S, Rees MI, Harvey RJ (2012) Mutations in the GlyT2 gene (*SLC6A5*) are a second major cause of startle disease. *J Biol Chem* 287: 28975-28985.

Chang Y, Weiss DS (1998) Substitutions of the highly conserved M2 leucine create spontaneously opening ρ1 γ-aminobutyric acid receptors. *Mol Pharmacol* 53: 511-523.

Chang Y, Weiss DS (1999) Allosteric activation mechanism of the α1β2γ2 γ-aminobutyric acid type A receptor revealed by mutation of the conserved M2 leucine. *Biophys J* 77: 2542-2551.

Charlier C, Coppieters W, Rollin F, Desmecht D, Agerholm JS, Cambisano N, Carta E, Dardano S, Dive M, Fasquelle C, Frennet JC, Hanset R, Hubin X, Jorgensen C, Karim L, Kent M, Harvey K, Pearce BR, Simon P, Tama N, Nie H, Vandeputte S, Lien S, Longeri M, Fredholm M, Harvey RJ, Georges M (2008) Highly effective SNP-based association mapping and management of recessive defects in livestock. *Nat Genet* 40: 449-454.

Chen JG, Liu-Chen S, Rudnick G (1997) External cysteine residues in the serotonin transporter. *Biochemistry* 36: 1479-1486.

Chen R, Wei H, Hill ER, Chen L, Jiang L, Han DD, Gu HH (2007) Direct evidence that two cysteines in the dopamine transporter form a disulfide bond. *Mol Cell Biochem* 298: 41-48.

Chung SK, Vanbellinghen JF, Mullins JG, Robinson A, Hantke J, Hammond CL, Gilbert DF, Freilinger M, Ryan M, Kruer MC, Masri A, Gurses C, Ferrie C, Harvey K, Shiang R, Christodoulou J, Andermann F, Andermann E, Thomas RH, Harvey RJ, Lynch JW, Rees MI (2010) Pathophysiological mechanisms of dominant and recessive *GLRA1* mutations in hyperekplexia. *J Neurosci* 30: 9612-9620.

Cohen PS, Zetter BR, Walsh ML (1972) Evidence that more deoxynucleotide kinase mRNA is transcribed than translated during T4 infection of *Escherichia coli*. *Virology* 49: 808-810.

Cohen SN, Chang AC, Hsu L (1972) Nonchromosomal antibiotic resistance in bacteria: genetic transformation of Escherichia coli by R-factor DNA. *Proc Natl Acad Sci USA* 69: 2110-2114.

Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J Am Chem Soc* 117: 5179-5197.

Coto E, Armenta D, Espinosa R, Argente J, Castro MG, Alvarez V (2005) Recessive hyperekplexia due to a new mutation (R100H) in the *GLRA1* gene. *Mov Disord* 20: 1626-1629.

Cui WW, Low SE, Hirata H, Saint-Amant L, Geisler R, Hume RI, Kuwada JY (2005) The zebrafish shocked gene encodes a glycine transporter and is essential for the function of early neural circuits in the CNS. *J Neurosci* 25: 6610-6620.

Cymes GD, Grosman C (2011) Tunable pKa values and the basis of opposite charge selectivities in nicotinic-type receptors. *Nature* 474: 526-530.

David-Watine B, Goblet C, de Saint Jan D, Fucile S, Devignot V, Bregestovski P, Korn H (1999) Cloning, expression and electrophysiological characterization of glycine receptor α subunit from zebrafish. *Neuroscience* 90: 303-17.

Davies JS, Chung S-K, Thomas RH, Robinson A, Hammond CL, Mullins JGL, Carta E, Pearce BR, Harvey K, Harvey RJ, Rees MI (2010) The glycinergic system in human startle disease: a genetic screening approach. *Front Mol Neurosci* 3: 8.

del Giudice EM, Coppola G, Bellini G, Cirillo G, Scuccimarra G, Pascotto A (2001) A mutation (V260M) in the middle of the M2 pore-lining domain of the glycine receptor causes hereditary hyperekplexia. *Eur J Hum Genet* 9: 873-876.

del Pino I, Paarmann I, Karas M, Kilimann MW, Betz H (2011) The trafficking proteins Vacuolar Protein Sorting 35 and Neurobeachin interact with the glycine receptor β-subunit. *Biochem Biophys Res Commun* 412: 435-440.

Doria Lamba L, Giribaldi G, De Negri E, Follo R, De Grandis E, Pintaudi M, Veneselli E (2007) A case of major form familial hyperekplexia: prenatal diagnosis and effective treatment with clonazepam. *J Child Neurol* 22: 769-772.

Downes GB, Granato M (2004) Acetylcholinesterase function is dispensable for sensory neurite growth but is critical for neuromuscular synapse stability. *Dev Biol* 270: 232-245.

Drapeau P, Saint-Amant L, Buss RR, Chong M, McDearmid JR, Brustein E (2002) Development of the locomotor network in zebrafish. *Prog Neurobiol* 68: 85-111.

Drögemüller C, Drögemüller M, Leeb T, Mascarello F, Testoni S, Rossi M, Gentile A, Damiani E, Sacchetto R (2008) Identification of a missense mutation in the bovine *ATP2A1* gene in congenital pseudomyotonia of Chianina cattle: an animal model of human Brody disease. *Genomics* 92: 474-477.

Dunbrack RL (2002) Rotamer libraries in the 21st century. *Curr Opin Struc Biol* 12: 431-440.

Dutertre S, Drwal M, Laube B, Betz H (2012) Probing the pharmacological properties of distinct subunit interfaces within heteromeric glycine receptors reveals a functional ββ agonist-binding site. *J Neurochem* 122: 38-47.

Eddy SR (2004) Where did the BLOSUM62 alignment score matrix come from? *Nat Biotech* 22: 1035-1036.

Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinf* 5: 113.

Elmslie FV, Hutchings SM, Spencer V, Curtis A, Covanis T, Gardiner RM, Rees M (1996) Analysis of *GLRA1* in hereditary and sporadic hyperekplexia: a novel mutation in a

family cosegregating for hyperekplexia and spastic paraparesis. *J Med Genet* 33: 435-436.

Eramian D, Eswar N, Shen MY, Sali A (2008) How well can the accuracy of comparative protein structure models be predicted? *Protein Sci* 17:1881-1893.

Escayg A, MacDonald BT, Meisler MH, Baulac S, Huberfeld G, An-Gourfinkel I, Brice A, LeGuern E, Moulard B, Chaigne D, Buresi C, Malafosse A (2000) Mutations of *SCN1A*, encoding a neuronal sodium channel, in two families with GEFS+. *Nat Genet* 24: 343-345.

Eswar N, Eramian D, Webb B, Shen M-Y, Sali A (2008) Protein structure modeling with MODELLER. *Methods Mol Biol* 426: 145-159.

Eswar N, John B, Mirkovic N, Fiser A, Ilyin VA, Pieper U, Stuart AC, Marti-Renom MA, Madhusudhan MS, Yerkovich B, Sali A (2003) Tools for comparative protein structure modeling and analysis. *Nucleic Acids Res* 31: 3375-3380.

Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen M-Y, Pieper U, Sali A (2006) Comparative protein structure modeling using Modeller. Curr Protoc Bioinformatics Chapter 5: Unit 5.6

Eulenburg V, Armsen W, Betz H, Gomeza J (2005) Glycine transporters: essential regulators of neurotransmission. *Trends Biochem Sci* 30: 325-333.

Fasquelle C, Sartelet A, Li W, Dive M, Tamma N, Michaux C, Druet T, Huijbers IJ, Isacke CM, Coppieters W, Georges M, Charlier C (2009) Balancing selection of a frame-shift mutation in the MRC2 gene accounts for the outbreak of the Crooked Tail Syndrome in Belgian Blue Cattle. *PLoS Genet* 5: e1000666.

Feng DF, Doolittle RF (1987) Progressive sequence alignment as a prerequisite to correct phylogenetic trees. *J Mol Evol* 25: 351-360.

Feng G (1998) Dual requirement for gephyrin in glycine receptor clustering and molybdoenzyme activity. *Science* 282: 1321-1324.

Flint AC, Liu X, Kriegstein AR (1998) Nonsynaptic glycine receptor activation during early neocortical development. *Neuron* 20: 43-53.

Forrest LR, Tavoulari S, Zhang YW, Rudnick G, Honig B (2007) Identification of a chloride ion binding site in Na^+/Cl^--dependent transporters. *Proc Natl Acad Sci* 104: 12761-12766.

Forsyth RJ, Gika AD, Ginjaar I, Tijssen MA (2007) A novel *GLRA1* mutation in a recessive hyperekplexia pedigree. *Mov Disord* 22: 1643-1645.

Fox JG, Averill DR, Hallett M, Schunk K (1984) Familial reflex myoclonus in Labrador Retrievers. *Am J Vet Res* 45: 2367-2370.

Friedrich T, Lambert AM, Masino MA, Downes GB (2012) Mutation of zebrafish dihydrolipoamide branched-chain transacylase E2 results in motor dysfunction and models maple syrup urine disease. *Dis Model Mech* 5: 248-258.

Gallivan JP, Dougherty DA (1999) Cation-π interactions in structural biology. *Proc Natl Acad Sci USA* 96: 9459-9464.

Geisler R, Rauch GJ, Geiger-Rudolph S, Albrecht A, van Bebber F, Berger A, Busch-Nentwich E, Dahm R, Dekens MP, Dooley C, Elli AF, Gehring I, Geiger H, Geisler M, Glaser S, Holley S, Huber M, Kerr A, Kirn A, Knirsch M, Konantz M, Küchler AM, Maderspacher F, Neuhauss SC, Nicolson T, Ober EA, Praeg E, Ray R, Rentzsch B, Rick JM, Rief E, Schauerte HE, Schepp CP, Schönberger U, Schonthaler HB, Seiler C, Sidi S, Söllner C, Wehner A, Weiler C, Nüsslein-Volhard C (2007) Large-scale mapping of mutations affecting zebrafish development. *BMC Genomics* 8: 11.

Gilbert SL, Ozdag F, Ulas UH, Dobyns WB, Lahn BT (2004) Hereditary hyperekplexia caused by novel mutations of *GLRA1* in Turkish families. *Mol Diagn* 8: 151-155.

Gill JL, Capper D, Vanbellinghen J-F, Chung S-K, Higgins RJ, Rees MI, Shelton GD, Harvey RJ (2011) Startle disease in Irish wolfhounds associated with a microdeletion in the glycine transporter GlyT2 gene. *Neurobiol Dis* 43: 184-189.

Giménez C, Pérez-Siles G, Martinez-Villarreal J, Arribas-González E, Jiméenez E, Núñez E, de Juan-Sanz J, Fernández-Sanchez E, Garcia-Tardón N, Ibáñez I, Romanelli V, Nevado J, James VM, Topf M, Chung SK, Thomas RH, Desviat LR, Aragón C, Zafra F, Rees MI, Lapunzina P, Harvey RJ, López-Corcuera B (2012) A novel dominant hyperekplexia mutation Y705C alters trafficking and biochemical properties of the presynaptic glycine transporter GlyT2. *J Biol Chem* 287: 28986-29002.

Gleason MR, Armisen R, Verdecia MA, Sirotkin H, Brehm P, Mandel G (2004) A mutation in *serca* underlies motility dysfunction in *accordion* zebrafish. *Dev Biol* 276: 441-451.

Gomeza J, Hülsmann S, Ohno K, Eulenburg V, Szöke K, Richter D, Betz H (2003) Inactivation of the glycine transporter 1 gene discloses vital role of glial glycine uptake in glycinergic inhibition. *Neuron* 40: 785-796.

Gomeza J, Ohno K, Hülsmann S, Armsen W, Eulenburg V, Richter DW, Laube B, Betz H (2003) Deletion of the mouse glycine transporter 2 results in a hyperekplexia phenotype and postnatal lethality. *Neuron* 40: 797-806.

Granato M, van Eeden FJ, Schach U, Trowe T, Brand M, Furutani-Seiki M, Haffter P, Hammerschmidt M, Heisenberg CP, Jiang YJ, Kane DA, Kelsh RN, Mullins MC, Odenthal J, Nüsslein-Volhard C (1996) Genes controlling and mediating locomotion behaviour of the zebrafish embryo and larva. *Development* 123: 399-413.

Gregory ML, Guzauskas GF, Edgar TS, Clarkson KB, Srivastava AK, Holden KR (2008) A novel *GLRA1* mutation associated with an atypical hyperekplexia phenotype. *J Child Neurol* 23: 1433-1438.

Grudzinska J, Schemm R, Haeger S, Nicke A, Schmalzing G, Betz H, Laube B (2005) The β subunit determines the ligand binding properties of synaptic glycine receptors. *Neuron* 45: 727-739.

Gundlach AL, Kortz G, Burazin TC, Madigan J, Higgins RJ (1993) Deficit of inhibitory glycine receptors in spinal cord from Peruvian Pasos: evidence for an equine form of inherited myoclonus. *Brain Res* 628: 263-270.

Haeger S, Kuzmin D, Detro-Dassen S, Lang N, Kilb M, Tsetlin V, Betz H, Laube B, Schmalzing G (2009) An intramembrane aromatic network determines pentameric assembly of Cys-loop receptors. *Nat Struct Mol Biol* 17: 90-98.

Harkin LA, Bowser DN, Dibbens LM, Singh R, Phillips F, Wallace RH, Richards MC, Williams DA, Mulley JC, Berkovic SF, Scheffer IE, Petrou S (2002) Truncation of the $GABA_A$-receptor γ2 subunit in a family with generalized epilepsy with febrile seizures plus. *Am J Hum Genet* 70: 530-536.

Harvey K, Duguid IC, Alldred MJ, Beatty SE, Ward H, Keep NH, Lingenfelter SE, Pearce BR, Lundgren J, Owen MJ, Smart TG, Lüscher B, Rees MI, Harvey RJ (2004) The GDP-GTP exchange factor collybistin: an essential determinant of neuronal gephyrin clustering. *J Neurosci* 24: 5816-5826.

Harvey RJ, Carta E, Pearce BR, Chung S-K, Supplisson S, Rees MI, Harvey K (2008) A critical role for glycine transporters in hyperexcitability disorders. *Front Mol Neurosci* 1: 1.

Harvey RJ, Depner UB, Wässle H, Ahmadi S, Heindl C, Reinold H, Smart TG, Harvey K, Schütz B, Abo-Salem OM, Zimmer A, Poisbeau P, Welzl H, Wolfer DP, Betz H, Zeilhofer HU, Müller U (2004) GlyR α3: An essential target for spinal PGE_2-mediated inflammatory pain sensitization. *Science* 304: 884-887.

Harvey RJ, Schmieden V, Holst Von A, Laube B, Rohrer H, Betz H (2000) Glycine receptors containing the α4 subunit in the embryonic sympathetic nervous system, spinal cord and male genital ridge. *Eur J Neurosci* 12: 994-1001.

Harvey RJ, Topf M, Harvey K, Rees MI (2008) The genetics of hyperekplexia: more than startle! *Trends Genet* 24: 439-447.

Harvey VL, Caley A, Müller UC, Harvey RJ, Dickenson AH (2009) A selective role for α3 subunit glycine receptors in inflammatory pain. *Front Mol Neurosci* 2: 14.

Heinze L, Harvey RJ, Haverkamp S, Wässle H (2006) Diversity of glycine receptors in the mouse retina: Localization of the α4 subunit. *J Comp Neurol* 500: 693-707.

Hibbs RE, Gouaux E (2011) Principles of activation and permeation in an anion-selective Cys-loop receptor. *Nature* 474: 54-60.

Higashijima S-I, Mandel G, Fetcho JR (2004) Distribution of protective glutamergic, glycinergic and GABAergic neurons in embryonic and larval zebrafish. *J Comp Neurol* 480: 1-18.

Hilf RJC, Dutzler R (2008) Structure of a potentially open state of a proton-activated pentameric ligand-gated ion channel. *Nature* 457: 115-119.

Hinsen K (2000) The molecular modeling toolkit: a new approach to molecular simulations. *J Comput Chem* 21: 79-85.

Hirata H, Carta E, Yamanaka I, Harvey RJ, Kuwada JY (2010) Defective glycinergic synaptic transmission in zebrafish motility mutants. *Front Mol Neurosci* 2: 26.

Hirata H, Saint-Amant L, Downes GB, Cui WW, Zhou W, Granato M, Kuwada JY (2005) Zebrafish *bandoneon* mutants display behavioral defects due to a mutation in the glycine receptor β-subunit. *Proc Natl Acad Sci* 102: 8345-8350.

Hirata H, Saint-Amant L, Waterbury J, Cui W, Zhou W, Li Q, Goldman D, Granato M, Kuwada JY (2004) *accordion*, a zebrafish behavioral mutant, has a muscle relaxation defect due to a mutation in the ATPase Ca^{2+} pump SERCA1. *Development* 131: 5457-5468.

Holland KD, Fleming MT, Cheek S, Moran JL, Beier DR, Meisler MH (2006) *De novo* exon duplication in a new allele of mouse *Glra1* (*spasmodic*). *Genetics* 174: 2245-2247.

Holman S, Morgan T, Baujat G, Cormier-Daire V, Cho TJ, Lees M, Samanich J, Tapon D, Hove H, Hing A, Hennekam R, Robertson S (2012) Osteopathia striata congenita with cranial sclerosis and intellectual disability due to contiguous gene deletions involving the WTX locus. *Clin Genet*, in press.

Hösl K, Reinold H, Harvey RJ, Müller U, Narumiya S, Zeilhofer HU (2006) Spinal prostaglandin E receptors of the EP_2 subtype and the glycine receptor α3 subunit, which mediate central inflammatory hyperalgesia, do not contribute to pain after peripheral nerve injury or formalin injection. *Pain* 126: 46-53.

Humeny A, Bonk T, Becker K, Jafari-Boroujerdi M, Stephani U, Reuter K, Becker CM (2002) A novel recessive hyperekplexia allele *GLRA1* (S231R): genotyping by MALDI-TOF mass spectrometry and functional characterisation as a determinant of cellular glycine receptor trafficking. *Eur J Hum Genet* 10: 188-196.

Hunter CA (1993) Sequence-dependent DNA structure. The role of base stacking interactions. *J Mol Biol* 230: 1025-1054.

Hurst JM, McMillan LE, Porter CT, Allen J, Fakorede A, Martin AC (2009) The SAAPdb web resource: a large-scale structural analysis of mutant proteins. *Hum Mutat* 30: 616-624.

Imboden M, Devignot V, Goblet C (2001) Phylogenetic relationships and chromosomal location of five distinct glycine receptor subunit genes in the teleost *Danio Rerio*. *Dev Genes Evol* 211: 415-422.

Islam R, Lynch JW (2012) Mechanism of action of the insecticides, lindane and fipronil, on glycine receptor chloride channels. *Br J Pharmacol* 165: 2707-2720.

Jakalian A, Bush BL, Jack DB, Bayly CI (2000) Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method. *J Comput Chem* 21: 132-146.

Jakalian A, Jack DB, Bayly CI (2002) Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J Comput Chem* 23: 1623-1641.

Jones DT (1999) GenTHREADER: an efficient and reliable protein fold recognition method for genomic sequences. *J Mol Biol* 287: 797-815.

Jones DT, Thornton JM (1996) Potential energy functions for threading. *Curr Opin Struc Biol* 6: 210-216.

Ju P, Aubrey KR, Vandenberg RJ (2004) Zn^{2+} inhibits glycine transport by glycine transporter subtype 1b. *J Biol Chem* 279: 22983-22991.

Kalscheuer VM, Musante L, Fang C, Hoffmann K, Fuchs C, Carta E, Deas E, Venkateswarlu K, Menzel C, Ullmann R, Tommerup N, Dalprà L, Tzschach A, Selicorni A, Lüscher B, Ropers HH, Harvey K, Harvey RJ (2010) A balanced chromosomal translocation disrupting *ARHGEF9* is associated with epilepsy, anxiety, aggression, and mental retardation. *Hum Mutat* 30: 61-68.

Kang HC, Jeong You S, Jae Chey M, Sam Baik J, Kim JW, Ki CS (2008) Identification of a *de nov*o Lys304Gln mutation in the glycine receptor α1 subunit gene in a Korean infant with hyperekplexia. *Mov Disord* 23: 610-613.

Kelsell DP, Norgett EE, Unsworth H, Teh MT, Cullup T, Mein CA, Dopping-Hepenstal PJ, Dale BA, Tadini G, Fleckman P, Stephens KG, Sybert VP, Mallory SB, North BV, Witt DR, Sprecher E, Taylor AE, Ilchyshyn A, Kennedy CT, Goodyear H, Moss C, Paige D, Harper JI, Young BD, Leigh IM, Eady RA, O'Toole EA (2005) Mutations in *ABCA12* underlie the severe congenital skin disease harlequin ichthyosis. *Am J Hum Genet* 76: 794-803.

Kim EY, Schrader N, Smolinsky B, Bedet C, Vannier C, Schwarz G, Schindelin H (2006) Deciphering the structural framework of glycine receptor anchoring by gephyrin. *EMBO J* 25: 1385-1395.

Kimura M, Taketani T, Horie A, Isumi H, Sejima H, Yamaguchi S (2006) Two Japanese families with hyperekplexia who have a Arg271Gln mutation in the glycine receptor α1 subunit gene. *Brain Dev* 28: 228-231.

Kingsmore SF, Giros B, Suh D, Bieniarz M, Caron MG, Seldin MF (1994) Glycine receptor β-subunit gene mutation in spastic mouse associated with LINE-1 element insertion. *Nat Genet* 7:136-141.

Kirsch J, Betz H (1998) Glycine-receptor activation is required for receptor clustering in spinal neurons. *Nature* 392: 717-720.

Kitchen DB, Decornez H, Furr JR, Bajorath J (2004) Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug Discov* 3: 935-949.

Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305: 567-580.

Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305: 567-580.

Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 157: 105-132.

Langosch D, Laube B, Rundström N, Schmieden V, Bormann J, Betz H (1994) Decreased agonist affinity and chloride conductance of mutant glycine receptors associated

with human hereditary hyperekplexia. *EMBO J* 13: 4223-4228.

Langosch D, Thomas L, Betz H (1988) Conserved quaternary structure of ligand-gated ion channels: the postsynaptic glycine receptor is a pentamer. *Proc Natl Acad Sci USA* 85: 7394-7398.

Lapunzina P, Sánchez JM, Cabrera M, Moreno A, Delicado A, de Torres ML, Mori AM, Quero J, Lopez Pajares I (2003) Hyperekplexia (startle disease): a novel mutation (S270T) in the M2 domain of the *GLRA1* gene and a molecular review of the disorder. *Mol Diagn* 7:125-128.

Lee CG, Kwon MJ, Yu HJ, Nam SH, Lee J, Ki CS, Lee M (2012) Clinical features and genetic analysis of children with hyperekplexia in Korea. *J Child Neurol*, in press.

Lefévre C, Audebert S, Jobard F, Bouadjar B, Lakhdar H, Boughdene-Stambouli O, Blanchet-Bardon C, Heilig R, Foglio M, Weissenbach J, Lathrop M, Prud'homme JF, Fischer J (2003) Mutations in the transporter *ABCA12* are associated with lamellar ichthyosis type 2. *Hum Mol Genet* 12: 2369-2378.

Lemke JR, Riesch E, Scheurenbrand T, Schubach M, Wilhelm C, Steiner I, Hansen J, Courage C, Gallati S, Bürki S, Strozzi S, Simonetti BG, Grunt S, Steinlin M, Alber M, Wolff M, Klopstock T, Prott EC, Lorenz R, Spaich C, Rona S, Lakshminarasimhan M, Kröll J, Dorn T, Krämer G, Synofzik M, Becker F, Weber YG, Lerche H, Böhm D, Biskup S (2012) Targeted next generation sequencing as a diagnostic tool in epileptic disorders. *Epilepsia* 53: 1387-1398.

Lévi S, Vannier C, Triller A (1998) Strychnine-sensitive stabilization of postsynaptic glycine receptor clusters. *J Cell Sci* 111: 335-345.

Lobley A, Sadowski MI, Jones DT (2009) pGenTHREADER and pDomTHREADER: new methods for improved protein fold recognition and superfamily discrimination. *Bioinformatics* 25: 1761-1767.

Lynagh T, Lynch JW (2012) Molecular mechanisms of Cys-loop ion channel receptor modulation by ivermectin. *Front Mol Neurosci* 5: 60.

Lynch JW (2009) Native glycine receptor subtypes and their physiological roles. *Neuropharmacol* 56: 303-309.

Mandel M, Higa A (1970) Calcium-dependent bacteriophage DNA infection. *J Mol Biol* 53: 159-162.

Mangin JM, Baloul M, Prado De Carvalho L, Rogister B, Rigo JM, Legendre P (2003) Kinetic properties of the α2 homo-oligomeric glycine receptor impairs a proper synaptic functioning. *J Physiol* 553: 369-386.

Manzke T, Niebert M, Koch UR, Caley A, Vogelgesang S, Hülsmann S, Ponimaskin E, Müller U, Smart TG, Harvey RJ, Richter DW (2010) Serotonin receptor 1A-modulated phosphorylation of glycine receptor α3 controls breathing in mice. *J Clin Invest* 120: 4118-4128.

Martin AC, Facchiano AM, Cuff AL, Hernandez-Boussard T, Olivier M, Hainaut P, Thornton JM (2002) Integrating mutation data and structural analysis of the TP53

tumor-suppressor protein. *Hum Mutat* 19: 149-164.

Masino MA, Fetcho JR (2005) Fictive swimming motor patterns in wild type and mutant larval zebrafish. *J Neurophys* 93: 3177-3188.

McGaughey GB, Gagné M, Rappé AK (1998) π-Stacking interactions. Alive and well in proteins. *J Biol Chem* 273: 15458-15463.

McGuffin LJ, Jones DT (2003) Improvement of the GenTHREADER method for genomic fold recognition. *Bioinformatics* 19: 874-881.

Meyer G, Kirsch J, Betz H, Langosch D (1995) Identification of a gephyrin binding motif on the glycine receptor β subunit. *Neuron* 15: 563-572.

Milani N, Mülhardt C, Weber RG, Lichter P, Kioschis P, Poustka A & Becker CM (1998) The human glycine receptor β subunit gene (*GLRB*): structure, refined chromosomal localization, and population polymorphism. Genomics 50: 341-345.

Mills JE, Dean PM (1996) Three-dimensional hydrogen-bond geometry and probability information from a crystal survey. *J Comput Aided Mol Des* 10: 607-622.

Miraglia Del Giudice E, Coppola G, Bellini G, Ledaal P, Hertz JM, Pascotto A (2003) A novel mutation (R218Q) at the boundary between the N-terminal and the first transmembrane domain of the glycine receptor in a case of sporadic hyperekplexia. *J Med Genet* 40: e71.

Mitchell SM, Lee E, Garcia ML, Stephan MM (2004) Structure and function of extracellular loop 4 of the serotonin transporter as revealed by cysteine-scanning mutagenesis. *J Biol Chem* 279: 24089-24099.

Molteni L, Dardano S, Parma P, Polli M, De Giovanni AM, Sironi G, Longeri M (2006) Ichthyosis in Chianina cattle. *Vet Rec* 158: 412-414.

Mongeon R, Gleason MR, Masino MA, Fetcho JR, Mandel G, Brehm P, Dallman JE (2008) Synaptic homeostasis in a zebrafish glial glycine transporter mutant. *J Neurophys* 100: 1716-1723.

Montgomery J, Wittwer CT, Palais R, Zhou L (2007) Simultaneous mutation scanning and genotyping by high-resolution DNA melting analysis. *Nat Protoc* 2: 59-66.

Mülhardt C, Fischer M, Gass P, Simon-Chazottes D, Guénet JL, Kuhse J, Betz H, Becker CM (1994) The spastic mouse: aberrant splicing of glycine receptor β subunit mRNA caused by intronic insertion of L1 element. *Neuron* 13: 1003-1015.

Nasevicius A, Ekker SC (2000) Effective targeted gene 'knockdown' in zebrafish. *Nat Genet* 26: 216-220.

Nasevicius A, Ekker SC (2001) The zebrafish as a novel system for functional genomics and therapeutic development applications. *Curr Opin Mol Ther* 3: 224-228.

Needleman SB, Wunsch CD (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol* 48: 443-453.

Nelson N (1998) The family of Na^+/Cl^- neurotransmitter transporters. *J Neurochem* 71: 1785-1803.

Ng PC (2003) SIFT: predicting amino acid changes that affect protein function. *Nucl Acids Res* 31: 3812-3814.

Nigro MA, Lim HCN (1992) Hyperekplexia and sudden neonatal death. *Pediatr Neurol* 8: 221-225.

Norregaard L, Frederiksen D, Nielsen EO, Gether U (1998) Delineation of an endogenous zinc-binding site in the human dopamine transporter. *EMBO J* 117: 4266-4273.

Notredame C, Higgins DG, Heringa J (2000) T-coffee: a novel method for fast and accurate multiple sequence alignment. *J Mol Biol* 302: 205-217.

Odenthal J, Haffter P, Vogelsang E, Brand M, van Eeden FJ, Furutani-Seiki M, Granato M, Hammerschmidt M, Heisenberg CP, Jiang YJ, Kane DA, Kelsh RN, Mullins MC, Warga RM, Allende ML, Weinberg ES, Nüsslein-Volhard C (1996) Mutations affecting the formation of the notochord in the zebrafish, *Danio rerio*. *Development* 123:103-15.

Odermatt A, Taschner PE, Khanna VK, Busch HF, Karpati G, Jablecki CK, Breuning MH, MacLennan DH (1996) Mutations in the gene encoding SERCA1, the fast-twitch skeletal muscle sarcoplasmic reticulum Ca^{2+} ATPase, are associated with Brody disease. *Nat Genet* 14: 191-194.

Orengo CA, Michie AD, Jones S, Jones DT, Swindells MB, Thornton JM (1997) CATH – a hierarchic classification of protein domain structures. *Structure* 5: 1093-1108.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera - A visualization system for exploratory research and analysis. *J Comput Chem* 25: 1605-1612.

Pierce K, Handford CA, Morris R, Vafa B, Dennis JA, Healy PJ, Schofield PR (2001) A nonsense mutation in the α1 subunit of the inhibitory glycine receptor associated with bovine myoclonus. *Mol Cell Neurosci* 17: 354-363.

Pierce KD, Handford CA, Morris R, Vafa B, Dennis JA, Healy PJ, Schofield PR (2001) A nonsense mutation in the α1 subunit of the inhibitory glycine receptor associated with bovine myoclonus. *Mol Cell Neurosci* 17: 354-363.

Pless SA, Hanek AP, Price KL, Lynch JW, Lester HA, Dougherty DA, Lummis SCR (2011) A cation-π interaction at a phenylalanine residue in the glycine receptor binding site is conserved for different agonists. *Mol Pharmacol* 79: 742-748.

Poon WT, Au KM, Chan YW, Chan KY, Chow CB, Tong SF, Lam CW (2006) Novel missense mutation (Y279S) in the *GLRA1* gene causing hyperekplexia. *Clin Chim Acta* 364:361-362.

Praveen V, Patole SK, Whitehall JS (2001) Hyperekplexia in neonates. *Postgrad Med J* 77: 570-572.

Pribilla I, Takagi T, Langosch D, Bormann J, Betz H (1992) The atypical M2 segment of the β subunit confers picrotoxinin resistance to inhibitory glycine receptor channels. *EMBO J* 11: 4305-4311.

Rees MI, Andrew M, Jawad S, Owen MJ (1994) Evidence for recessive as well as dominant forms of startle disease (hyperekplexia) caused by mutations in the α1 subunit of the inhibitory glycine receptor. *Hum Mol Genet* 3: 2175-2179.

Rees MI, Harvey K, Pearce BR, Chung S-K, Duguid IC, Thomas P, Beatty S, Graham GE, Armstrong L, Shiang R, Abbott KJ, Zuberi SM, Stephenson JBP, Owen MJ, Tijssen MAJ, van den Maagdenberg AMJM, Smart TG, Supplisson S, Harvey RJ (2006) Mutations in the gene encoding GlyT2 (*SLC6A5*) define a presynaptic component of human startle disease. *Nat Genet* 38: 801-806.

Rees MI, Harvey K, Ward H, White JH, Evans L, Duguid IC, Hsu CC, Coleman SL, Miller J, Baer K, Waldvogel HJ, Gibbon F, Smart TG, Owen MJ, Harvey RJ, Snell RG (2003) Isoform heterogeneity of the human gephyrin gene (*GPHN*), binding domains to the glycine receptor, and mutation analysis in hyperekplexia. *J Biol Chem* 278: 24688-24696.

Rees MI, Lewis TM, Kwok JBJ, Mortier GR, Govaert P, Snell RG, Schofield PR, Owen MJ (2002) Hyperekplexia associated with compound heterozygote mutations in the β-subunit of the human inhibitory glycine receptor (*GLRB*). *Hum Mol Genet* 11: 853-860.

Rees MI, Lewis TM, Vafa B, Ferrie C, Corry P, Muntoni F, Jungbluth H, Stephenson JB, Kerr M, Snell RG, Schofield PR, Owen MJ (2001) Compound heterozygosity and nonsense mutations in the α1-subunit of the inhibitory glycine receptor in hyperekplexia. Hum Genet 109: 267-270.

Reiss J, Gross-Hardt S, Christensen E, Schmidt P, Mendel RR, Schwarz, G (2001) A mutation in the gene for the neurotransmitter receptor-clustering protein gephyrin causes a novel form of molybdenum cofactor deficiency. *Am J Hum Genet* 68: 208-213.

Reiss J, Hahnewald R (2011) Molybdenum cofactor deficiency: Mutations in *GPHN*, *MOCS1*, and *MOCS2*. *Hum Mutat* 32: 10-18.

Roux MJ, Supplisson S (2000) Neuronal and glial glycine transporters have different stoichiometries. *Neuron* 25: 373-383.

Ryan SG, Buckwalter MS, Lynch JW, Handford CA, Segura L, Shiang R, Wasmuth JJ, Camper SA, Schofield P, O'Connell P (1994) A missense mutation in the gene encoding the α1 subunit of the inhibitory glycine receptor in the *spasmodic* mouse. *Nat Genet* 7: 131-135.

Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234: 779-815.

Sartelet A, Klingbeil P, Franklin CK, Fasquelle C, Géron S, Isacke CM, Georges M, Charlier C (2012) Allelic heterogeneity of Crooked Tail Syndrome: result of balancing selection? *Anim Genet*, in press.

Saul B, Kuner T, Sobetzko D, Brune W, Hanefeld F, Meinck HM, Becker CM (1999) Novel *GLRA1* missense mutation (P250T) in dominant hyperekplexia defines an intracellular determinant of glycine receptor channel gating. *J Neurosci* 19: 869-877.

Scain AL, Le Corronc H, Allain AE, Muller E, Rigo JM, Meyrand P, Branchereau P, Legendre P (2010) Glycine release from radial cells modulates the spontaneous activity and its propagation during early spinal cord development. *J Neurosci* 30: 390-403.

Schneider VA, Granato M (2006) The myotomal diwanka (lh3) glycosyltransferase and type XVIII collagen are critical for motor growth cone migration. *Neuron* 50: 683-695.

Schorderet DF, Pescia G, Bernasconi A, Regli F (1994) An additional family with startle disease and a G1192A mutation at the α1 subunit of the inhibitory glycine receptor gene. *Hum Mol Genet* 3: 1201.

Seri M, Bolino A, Galietta LJ, Lerone M, Silengo M, Romeo G (1997) Startle disease in an Italian family by mutation (K276E): the alpha-subunit of the inhibiting glycine receptor. *Hum Mutat* 9: 185-187.

Shan Q, Han L & Lynch JW (2011) β Subunit M2-M3 loop conformational changes are uncoupled from α1 β glycine receptor channel gating: implications for human hereditary hyperekplexia. *PLoS One*. 6: e28105.

Sharma SD, Sarna A, Mukhopadhyay S (2006) Neonatal hyperekplexia: the stiff-baby syndrome. *Indian Pediatr* 43: 539-541.

Shen MY, Sali A (2006) Statistical potential for assessment and prediction of protein structures. *Protein Sci* 15: 2507-2524.

Shiang R, Ryan SG, Zhu YZ, Fielder TJ, Allen RJ, Fryer A, Yamashita S, O'Connell P, Wasmuth JJ (1995) Mutational analysis of familial and sporadic hyperekplexia. *Ann Neurol* 38: 85-91.

Shiang R, Ryan SG, Zhu YZ, Hahn AF, O'Connell P, Wasmuth JJ (1993) Mutations in the α1 subunit of the inhibitory glycine receptor cause the dominant neurologic disorder, hyperekplexia. *Nat Genet* 5: 351-358.

Shuster DE, Kehrli ME Jr, Ackermann MR, Gilbert RO (1992) Identification and prevalence of a genetic defect that causes leukocyte adhesion deficiency in Holstein cattle. *Proc Natl Acad Sci USA* 89: 9225-9229.

Simon J, Wakimoto H, Fujita N, Lalande M, Barnard EA (2004) Analysis of the set of $GABA_A$ receptor genes in the human genome. *J Biol Chem* 279: 41422-41435.

Singh SK, Piscitelli CL, Yamashita A, Gouaux E (2008) A competitive inhibitor traps LeuT in an open-to-out conformation. *Science* 322: 1655-1661.

Sirén A, Legros B, Chahine L, Misson J-P, Pandolfo M (2006) Hyperekplexia in Kurdish families: a possible *GLRA1* founder mutation. *Neurology* 67: 137-139.

Smith TF, Waterman MS (1981) Identification of common molecular subsequences. *J Mol Biol* 147: 195-197.

Söding J, Biegert A, Lupas AN (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* 33: W244-W248.

Supplisson S, Roux MJ. Why glycine transporters have different stoichiometries (2002) *FEBS Lett* 529: 93-101.

Tarpey PS, Smith R, Pleasance E, Whibley A, Edkins S, Hardy C, O'Meara S, Latimer C (2009) A systematic, large-scale resequencing screen of X-chromosome coding exons in mental retardation. *Nat Genet* 41: 535-543.

Tennessen JA, Bigham AW, O'Connor TD, Fu W, Kenny EE, Gravel S, McGee S, Do R, Liu X, Jun G, Kang HM, Jordan D, Leal SM, Gabriel S, Rieder MJ, Abecasis G, Altshuler D, Nickerson DA, Boerwinkle E, Sunyaev S, Bustamante CD, Bamshad MJ, Akey JM; Broad GO; Seattle GO; NHLBI Exome Sequencing Project (2012) Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science* 337: 64-69.

Thomas AC, Cullup T, Norgett EE, Hill T, Barton S, Dale BA, Sprecher E, Sheridan E, Taylor AE, Wilroy RS, DeLozier C, Burrows N, Goodyear H, Fleckman P, Stephens KG, Mehta L, Watson RM, Graham R, Wolf R, Slavotinek A, Martin M, Bourn D, Mein CA, O'Toole EA, Kelsell DP (2006) *ABCA12* is the major harlequin ichthyosis gene. *J Invest Dermatol* 126: 2408-2413.

Thomas RH, Stephenson JBP, Harvey RJ, Rees MI (2010) Hyperekplexia: Stiffness, startle and syncope. *J Pediatr Neurol* 8: 11-14.

Thompson AJ, Lester HA, Lummis SCR (2010) The structural basis of function in Cys-loop receptors. *Quart Rev Biophys* 43: 449-499.

Thomsen B, Horn P, Panitz F, Bendixen E, Petersen AH, Holm LE, Nielsen VH, Agerholm JS, Arnbjerg J, Bendixen C (2006) A missense mutation in the bovine *SLC35A3* gene, encoding a UDP-N-acetylglucosamine transporter, causes complex vertebral malformation. *Genome Res* 16: 97-105.

Tijssen MA, Shiang R, van Deutekom J, Boerman RH, Wasmuth JJ, Sandkuijl LA, Frants RR, Padberg GW (1995) Molecular genetic re-evaluation of the Dutch hyperekplexia family. *Arch Neurol* 52: 578-582.

Traka M, Seburn KL, Popko B (2006) *Nmf11* is a novel ENU-induced mutation in the mouse glycine receptor α1 subunit. *Mamm Genome* 17: 950-955.

Tsai CH, Chang FC, Su YC, Tsai FJ, Lu MK, Lee CC, Kuo CC, Yang YW, Lu CS (2004) Two novel mutations of the glycine receptor gene in a Taiwanese hyperekplexia family. *Neurology* 63: 893-896.

Unwin N (1993) Nicotinic acetylcholine receptor at 9Å resolution. *J Mol Biol* 229: 1101-1101.

Unwin N (2005) Refined structure of the nicotinic acetylcholine receptor at 4Å Resolution. *J Mol Biol* 346: 967-989.

Vandeputte S, Brihoum M, Hubin X, Georges M, Charlier C, Desmecht D, Guyot H, Rollin F (2006) Congenital muscular dystonia (CMD): a new congenital pathology in

Belgian Blue calves. In Proceedings of the 24th World Buiatrics Congress, Nice, 15-19th October 2006 (ref. OS05-2).

Vattemi G, Gualandi F, Oosterhof A, Marini M, Tonin P, Rimessi P, Neri M, Guglielmi V, Russignan A, Poli C, van Kuppevelt TH, Ferlini A, Tomelleri G (2010) Brody disease: insights into biochemical features of *SERCA1* and identification of a novel mutation. *J Neuropathol Exp Neurol* 69: 246-252.

Vergouwe MN, Tijssen MA, Peters AC, Wielaard R, Frants RR (1999) Hyperekplexia phenotype due to compound heterozygosity for *GLRA1* gene mutations. *Ann Neurol* 46:634-638.

Vigevano F, Di Capua M, Dalla Bernardina B (1989) Startle disease: an avoidable cause of sudden infant death. *Lancet* 1: 216.

Villmann C, Oertel J, Ma-Högemeier ZL, Hollmann M, Sprengel R, Becker K, Breitinger HG, Becker CM. Functional complementation of *Glra1*$^{spd\text{-}ot}$, a glycine receptor subunit mutant, by independently expressed C-terminal domains. *J Neurosci* 2009 29: 2440-2452.

Wada A (1976) The alpha-helix as an electric macro-dipole. *Adv Biophys* 9: 1-63.

Wallace RH, Wang DW, Singh R, Scheffer IE, George AL Jr, Phillips HA, Saar K, Reis A, Johnson EW, Sutherland GR, Berkovic SF, Mulley JC (1998) Febrile seizures and generalized epilepsy associated with a mutation in the Na^+-channel β1 subunit gene *SCN1B*. *Nat Genet* 19: 366-370.

Wang J, Wang W, Kollman PA, Case DA (2006) Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graph Model* 25: 247-260.

Wang M, Wen H, Brehm P (2008) Function of neuromuscular synapses in the zebrafish choline-acetyltransferase mutant *bajan*. *J Neurophysiol* 100: 1995-2004.

Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res* 35: W407-W410.

Wittwer CT (2009) High-resolution DNA melting analysis: advancements and limitations. *Hum Mutat* 30: 857-859.

Yamashita A, Singh SK, Kawate T, Jin Y, Gouaux E (2005) Crystal structure of a bacterial homologue of Na^+/Cl^--dependent neurotransmitter transporters. *Nature* 437: 215-223.

Yang Z, Taran E, Webb TI, Lynch JW (2012) Stoichiometry and subunit arrangement of α1β glycine receptors as determined by atomic force microscopy. *Biochemistry*, in press.

Young TL, Cepko CL (2004) A role for ligand-gated ion channels in rod photoreceptor development. *Neuron* 41: 867-879.

Young-Pearse TL, Ivic L, Kriegstein AR, Cepko CL (2006) Characterization of mice with targeted deletion of glycine receptor α2. *Mol Cell Biol* 26: 5728-5734.

Zeilhofer HU, Studler B, Arabadzisz D, Schweizer C, Ahmadi S, Layh B, Bösl MR, Fritschy JM (2005) Glycinergic neurons expressing enhanced green fluorescent protein in bacterial artificial chromosome transgenic mice. *J Comp Neurol* 482: 123-141.

Zeller J, Granato M (1999) The zebrafish *diwanka* gene controls an early step of motor growth cone migration. *Development* 126: 3461-3472.

Zomot E, Bendahan A, Quick M, Zhao Y, Javitch JA, Kanner BI (2007) Mechanism of chloride interaction with neurotransmitter:sodium symporters. *Nature* 449: 726-730.

Zomot E, Kanner BI (2003) The interaction of the γ-aminobutyric acid transporter GAT-1 with the neurotransmitter is selectively impaired by sulfhydryl modification of a conformationally sensitive cysteine residue engineered into extracellular loop IV. *J Biol Chem* 278: 42950-42958.

Zoons E, Ginjaar IB, Bouma PA, Carpay JA, Tijssen MA (2012). A new hyperekplexia family with a recessive frameshift mutation in the *GLRA1* gene. *Mov Disord* 27:795-796.

APPENDICES

cDNA position	Exon	Protein Precursor/mature	Type	dbSNP ID	SIFT	Polyphen-2	Overall Prediction
c.A54G	2	p.E18E / p.E-5E	Synonymous	rs11552656	-	-	Benign
c.T121C	2	p.S41P/p.S19P	Missense	rs139168167	Tolerated (0.14)	Possibly damaging (0.940)	Benign
c.G417T	5	p.W139C/p.W117C	Missense	rs62636642	Damaging (0)	Probably damaging (1.000)	Benign*
c.T423C	5	p.P141P / p.P119P	Synonymous	-	-	-	Benign
c.C822T	8	p.Y274Y / p.Y252Y	Synonymous	-	-	-	Benign
c.T831C	8	p.T277T / p.T255T	Synonymous	rs12507409	-	-	Benign*
c.DA1072	9	p.R359GfsX26 / p.R347GfsX26	Deletion / Frameshift	rs34766493	-	-	Likely pathogenic
c.T948C	9	p.L316L / p.L294L	Synonymous	rs1801154	-	-	Benign*
c.A1044G	9	p.A348A / p.A326A	Synonymous	-	-	-	Benign
c.C1385G	10	p.A462G / p.A440G	Missense	-	Tolerated (0.31)	Benign (0.051)	Benign

Appendix Table 1. Single nucleotide polymorphisms in *GLRB* found in 1000 Genomes project release 20100804
*Population data in dbSNP predicts that these variants are unlikely to be disease causing.
1000 Genomes Browser: http://browser.1000genomes.org/index.html
dbSNP: http://www.ncbi.nlm.nih.gov/projects/SNP/;
SIFT: http://sift.jcvi.org/www/SIFT_enst_submit.html with ENSP00000264428
Polyphen-2: http://genetics.bwh.harvard.edu/pph2/index.shtml

cDNA position	Exon	Protein Precursor/mature	Type	dbSNP ID or reference	SIFT	Polyphen-2	Overall Prediction
c.G22C	1	p.A8P / p.A-15P	Missense	rs139961228	Tolerated (0.22)	Benign (0.000)	Benign
c.G92A	2	p.K31R / p.K9R	Missense	rs149863285	Tolerated (0.22)	Possibly damaging (0.956)	Benign
c.T121C	2	p.S41P / p.S19P	Missense	rs139168167	Tolerated (0.14)	Possibly damaging (0.940)	Benign
c.C146G	3	p.A49G / p.A27G	Missense	rs141963848	Damaging (0.02)	Benign (0.071)	Benign
c.G185A	3	p.R62K / p.R40K	Missense	rs144279427	Tolerated (1)	Benign (0.001)	Benign
c.A404G	5	p.Y135C / p.Y113C	Missense	rs141265285	Damaging (0)	Benign (0.331)	Pathogenic?
c.T415C	5	p.W139R / p.W177R	Missense	rs143539995	Damaging (0)	Probably damaging (0.999)	Pathogenic?
c.G417T	5	p.W139C / p.W117C	Missense	rs62636642	Damaging (0)	Probably damaging (1.000)	Benign*
c.A446G	5	p.K149R / p.K127R	Missense	rs140878124	Damaging (0)	Probably damaging (0.997)	Pathogenic?
c.C499T	5	p.R167C / p.R145C	Missense	rs145671356	Damaging (0.01)	Probably damaging (1.000)	Pathogenic?
c.G511A	5	p.V171I / p.V149I	Missense	rs138787100	Tolerated (0.22)	Benign (0.049)	Benign
c.G587A	7	p.R196H / p.R174H	Missense	rs141976916	Damaging (0.05)	Possibly damaging (0.910)	Pathogenic?
c.G653A	7	p.G218E / p.G196E	Missense	rs146321769	Tolerated (0.15)	Benign (0.049)	Benign
c.A676G	7	p.I226V / p.I204V	Missense	rs138300557	Tolerated (0.47)	Benign (0.004)	Benign
c.C688T	7	p.Q230X / p.Q208X	Nonsense	rs151023274	-	-	Pathogenic?
c.G752A	7	p.G251D / p.G229D	Missense	rs121909749	Damaging (0)	Probably damaging (1.0)	Pathogenic?
c.G766A	8	p.V256M / p.V234M	Missense	rs144334540	Damaging (0)	Probably damaging (0.999)	Pathogenic?
c.G799A	8	p.G267S / p.G245S	Missense	rs142433300	Damaging (0.02)	Probably damaging (1.0)	Pathogenic?
c.G967C	9	p.V323L / p.V301L	Missense	rs147779008	Damaging (0)	Possibly damaging (0.920)	Pathogenic?
c.DA1072	9	p.R359GfsX26 / p.R347GfsX26	Deletion / Frameshift	rs34766493	-	-	Pathogenic?
c.G1133A	10	p.G378E / p.G356E	Missense	rs149915285	Tolerated (0.85)	Benign (0.010)	Benign
c.C1385G	10	p.A462G / p.A440G	Missense	rs148031091	Tolerated (0.22)	Benign (0.051)	Benign

Appendix Table 2. Single nucleotide polymorphisms in *GLRB* found in in NCBI dbSNP Build 136 resulting in coding non-synonymous changes, frameshifts or deletions. *Population data in dbSNP predicts that these variants are unlikely to be disease causing.
dbSNP: http://www.ncbi.nlm.nih.gov/projects/SNP/
SIFT: http://sift.jcvi.org/www/SIFT_enst_submit.html with ENSP00000264428
Polyphen-2: http://genetics.bwh.harvard.edu/pph2/index.shtml

Primer	Forward and reverse primers	Amplicon
glrbb ex1	5'-GTGCTTGTGATTTAGTTTCTTAACG-3' 5'-CCATACAGAAGCAGTGTTTAATGAC-3'	250 bp
glrbb ex2	5'-GAATCTTTAGCATGGTTTCATGTAGT-3' 5'-ACAGTGTTCAGAGTGGTATGTATTG-3'	239 bp
glrbb ex3	5'-AAATCTTTCGTTTGATAGATGGGATTG-3' 5'-GGGTGATGCTGAATGTGAATGA-3'	164 bp
glrbb ex4	5'-GCTCATCTTCAAGGTAGACAGT-3' 5'-CACAAAGACAGAGGTGTGGT-3'	328 bp
glrbb ex5	5'-TCACTGTCTGGTTCTCAGATGT-3' 5'-CCCTTTCCCATTTCATTTATCTTCTTTAT-3'	196 bp
glrbb ex6	5'-TTCATTTAAGAGCCAGCATATATTATTGTT-3' 5'-CACAGAGGAAGCTTATACACACT-3'	250 bp
glrbb ex7	5'-ATCATAGGAGGGCTAATAATTTTGACTT-3' 5'-TGATGAGGTGCTGATGCTTG-3'	261 bp
glrbb ex8	5'-TGTGTAAGCTCATCTCTCCTCAG-3' 5'-TGGTGAGCATAAATAACACCACAA-3'	411 bp
glrbb ex9	5'-GCAGAGATCTGACATGTTTTCTCTCTAT-3' 5'-CCGAGACGATTGGAATGAGGGA-3'	425 bp

Appendix Table 3 – Primers used for amplification and DNA sequencing of zebrafish *glrbb*.

Primer	Forward and reverse primers	Amplicon
slc6a5 ex1	5'-TTCCCACGGAGATATAGAAAGC-3' 5'-AACACACAAAGCGTAAAGTAATGA-3'	151 bp
slc6a5 ex2	5'-ACAGCTACTGCCTACATCTAATGTG-3' 5'-TGGCTGATTCTGCCGGATTG-3'	537 bp
slc6a5 ex3	5'-ATGAGTTGTAATAGTCTGTAATAAACGAA-3' 5'-CTTGTGTACTGCCTCTTGAA-3'	264 bp
slc6a5 ex4	5'-TTTATTTCCTTCATCGGCACTTT-3' 5'-GCTTAACGCAGTTTAGCCAT-3'	218 bp
slc6a5 ex5	5'-GCTGTAAGTCTGCTTTAAGTGTAATTTCTA-3' 5'-TGTGTGTGTATGAGAGAGAGAGAGG-3'	283 bp
slc6a5 ex6	5'-AAGTTGATTCCAAGTTCTGTCAT-3' 5'-CTGCAGATAAAGCACACAGATAG-3'	232 bp
slc6a5 ex7	5'-CTGCAGGATATTTGGATGTTAATCAC-3' 5'-CTTTAGAAACAGCTGACATCAGGAG-3'	226 bp
slc6a5 ex8	5'-CCTTAATAAGTTTGGTTTATTGTGATGTC-3' 5'-CAGAAAGCATTATACCATCAAACATAGG-3'	245 bp
slc6a5 ex9	5'-GTATTTCTTTCTCTCTTTCTCTGCT-3' 5'-AACAGCTTTTATTGAAATACAAATTGTCTA-3'	192 bp
slc6a5 ex10	5'-GTAATGTATCCTGTGCTTTGTCT-3' 5'-AATACAAAACACATGCCATTACAC-3'	198 bp
slc6a5 ex11	5'-TTTCATTGTCTTCCAAGTATGAGAACTC-3' 5'-GTAAGTAACATAAGGAAGAAGATGATGGC-3'	150 bp
slc6a5 ex12	5'-ATTTGCCCAGCTGTATCAAATAAAT-3' 5'-ATTAAAGTAAGTGTTGGTAGTTGAGAAT-3'	230 bp
slc6a5 ex13	5'-TAAAGTGTGTCTATCTGTGTGTGTT-3' 5'-AGCATCTGCTAAATGACTAAATGTAAAT-3'	195 bp
slc6a5 ex14	5'-ACATGTTTGCTCATTTATGTTATTTCT-3' 5'-ATGTATCTACAGGAATGGTCGT-3'	199 bp
slc6a5 ex15	5'-GCATATTAATCAGTATTTCTCTCTCTGTG-3' 5'-GTAACTGTATAGTAGAGAACAGAAAAAAAC-3'	248 bp
slc6a5 ex16	5'-GTCCCGGCTTTATGCACTATTC-3' 5'-CTCGGGATATTTCACACTCACAG-3'	243 bp

Appendix Table 4 – Primers used for amplification and DNA sequencing of zebrafish *slc6a5*.

PUBLICATIONS AND PRESENTATIONS

Some of the work described in this thesis has been published in the following peer-reviewed articles/reviews, or presented at internal and external scientific meetings.

PUBLICATIONS

James VM, Gill JL, Topf M, Harvey RJ (2012) Molecular mechanisms of glycine transporter GlyT2 mutations in startle disease. *Biol Chem* 393: 283-289.

Gill JL*, James VM*, Carta E, Harris D, Topf M, Scholes SF, Hateley G, Harvey RJ (2012) Identification of congenital muscular dystonia 2 associated with an inherited GlyT2 defect in Belgian Blue cattle from the United Kingdom. *Anim Genet* 43: 267-270. *Equal contribution.

Carta E*, Chung SK*, James VM*, Robinson A, Gill JL, Remy N, Vanbellinghen JF, Drew CJ, Cagdas S, Cameron D, Cowan FM, Del Toro M, Graham GE, Manzur AY, Masri A, Rivera S, Scalais E, Shiang R, Sinclair K, Stuart CA, Tijssen MA, Wise G, Zuberi SM, Harvey K, Pearce BR, Topf M, Thomas RH, Supplisson S, Rees MI, Harvey RJ (2012) Mutations in the GlyT2 gene (*SLC6A5*) are a second major cause of startle disease. *J Biol Chem* 287: 28975-28985. *Equal contribution.

Giménez C, Pérez-Siles G, Martínez-Villarreal J, Arribas-González E, Jiménez E, Núñez E, de Juan-Sanz J, Fernández-Sánchez E, García-Tardón N, Ibáñez I, Romanelli V, Nevado J, James VM, Topf M, Chung SK, Thomas RH, Desviat LR, Aragón C, Zafra F, Rees MI, Lapunzina P, Harvey RJ, López-Corcuera B (2012) A novel dominant hyperekplexia mutation Y705C alters trafficking and biochemical properties of the presynaptic glycine transporter GlyT2. *J Biol Chem* 287: 28986-29002.

James VM, Bode, A, Gill JL, Nielsen M, Cowan FM, Vujic M, Thomas RH, Rees MI, Harvey K, Keramidas A, Topf M, Ginjarr I, Lynch JW, Harvey RJ (2012) Novel missense mutation in the glycine receptor β subunit gene (*GLRB*) in startle disease. *Neurobiol Dis*, submitted.

PRIZES AND AWARDS

Winner of the April 2012 UCL School of Pharmacy Ph.D. research day prize for best speaker.

POSTERS AND PRESENTATIONS

James VM, Carta E, Topf M, Harvey RJ. Molecular insights into disease mechanisms in glycinergic disorders. School of Pharmacy Neuroscience Retreat, 12th-13th July 2010, Robinson College, Cambridge. *Poster presentation*.

James VM, Carta E, Ganser LR, Dallman JE, Topf M, Harvey RJ. Defective glycinergic synaptic transmission in zebrafish motility mutants. Society for Neuroscience Annual

Meeting, 13th - 17th November 2010, San Diego Convention Center, San Diego, CA, USA. *Poster presentation*.

James VM, Carta E, Chung S-K, Robinson A, Gill JL, Remy N, Vanbellinghen J-F, Harvey K, Pearce BR, Topf M, Thomas R, Rees MI, Harvey RJ. Mutations in the GlyT2 gene (*SLC6A5*) are the second major cause of hyperekplexia. School of Pharmacy Ph.D. Research Day, 15th April 2011, School of Pharmacy, London. *Poster presentation*.

James VM. Mutations in the GlyT2 gene (*SLC6A5*) are the second major cause of hyperekplexia. Bloomsbury Colleges Research Students Symposium, 26th May 2011, The Royal Veterinary College, Hawkshead campus, Hertfordshire. *Oral presentation*.

James VM, Carta E, Chung S-K, Robinson A, Gill JL, Remy N, Vanbellinghen J-F, Harvey K, Pearce BR, Topf M, Thomas R, Rees MI, Harvey RJ. Mutations in the GlyT2 gene are the second major cause of hyperekplexia. The Institute of Structural and Molecular Biology Retreat, 21 - 22nd June 2011, Robinson College, Cambridge. *Poster presentation*.

James VM, Bode A, Ginjarr I, Nielsen M, Gill JL, Thomas RH, Rees MI, Topf M, Lynch JW, Harvey RJ. Novel mutations in the glycine receptor β subunit gene (*GLRB*) in startle disease. Genomic Disorders 2012: The Genomics of Rare Diseases, 21st - 24th March 2012, Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK. *Poster presentation*.

Harvey RJ, Carta E, Chung SK, James VM, Vanbellinghen JF, Pearce BR, Topf M, Thomas RH, Supplisson S, Rees MI. Mutations in the GlyT2 gene (*SLC6A5*) are a second major cause of hyperekplexia. Genomic Disorders 2012: The Genomics of Rare Diseases, 21st - 24th March 2012, Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK. *Poster presentation*.

James VM. Novel mutations in the glycine receptor β subunit gene (*GLRB*) in startle disease. UCL School of Pharmacy Ph.D. Research Day, 2nd April 2012, UCL School of Pharmacy, London. *Oral presentation*.

James VM, Bode A, Ginjarr I, Nielsen M, Gill JL, Thomas RH, Rees MI, Topf M, Lynch JW, Harvey RJ. Novel mutations in the glycine receptor β subunit gene (*GLRB*) in startle disease. UCL School of Pharmacy Neuroscience Retreat, 6th July 2012, ZSL London Zoo, London. *Poster presentation*.

James VM, Bode A, Ginjarr I, Nielsen M, Gill JL, Thomas RH, Rees MI, Topf M, Lynch JW, Harvey RJ. Novel mutations in the glycine receptor β subunit gene (*GLRB*) in startle disease. Bloomsbury Colleges Research Student's Symposium, 16th July 2012, The Royal Veterinary College, Camden, London. *Poster presentation*.

James VM, Bode A, Ginjarr I, Nielsen M, Gill JL, Thomas RH, Rees MI, Lynch JW, Harvey RJ, Topf M. Novel mutations in the glycine receptor β subunit gene (*GLRB*) in startle disease. 20th Annual International Conference on Intelligent Systems for Molecular Biology, 14th - 17th July 2012, Long Beach Convention & Entertainment Center, Long Beach, CA, USA. *Poster presentation*.

LIST OF ABBREVIATIONS

AD	Autosomal dominant
Amp	Ampicillin
AR	Autosomal recessive
ARHGEF9	Cdc42 guanine nucleotide exchange factor 9 gene
BCAA	Branched-chain amino acids
BCKD	α-keto acid dehydrogenase enzyme complex
beo	*Bandoneon* mutant
BLAD	Bovine leucocyte adhesion deficiency
BLAT	Basic Local Alignment Tool
CH	Compound heterozygosity
CMD	Congenital muscular dystonia
CMD1	Congenital muscular dystonia type 1
CMD2	Congenital muscular dystonia type 2
CNS	Central nervous system
DAT	Dopamine transporter
dNTPs	Deoxyribonucleotides
DNA	Deoxyribonucleic acid
DOPE	Discrete Optimised Potential Energy
EB	Elution buffer
ECD	Extracellular domain
EDTA	Ethylenediaminetetraacetic acid
EGFP	Enhanced green fluorescent protein
EL	Extracellular loop
ELIC	*Erwinia chrysanthemi* ligand-gated ion channel
ENU	Ethylnitrosourea
GABA	γ-aminobutyric acid
GAT	GABA transporter
GCS	Glycine cleavage system
GEF	Guanine nucleotide exchange factor
GFP	Green fluorescent protein
GLIC	*Gloeobacter violaceus* ligand-gated ion channel
GLRA1	Glycine receptor α1 subunit gene
GLRB	Glycine receptor β subunit gene
GluCl	Glutamate-gated chloride channel
GlyR	Glycine receptor
GlyT	Glycine transporter

GPHN	Gephyrin gene
HEK293	Human Embryonic Kidney 293
hpf	Hours post fertilisation
ICM	Inherited congenital myoclonus
IRES	Internal ribosome entry site
LB	Luria-Bertani
Leu	Leucine
LeuT	Leucine transporter
MO	Morpholino
MSUD	Maple syrup urine disease
nAChR	Nicotinic acetylcholine receptor
NET	Noradrenaline transporter
NMDA	N-methyl-D-aspartate
NMDAR	NMDA receptor
OMIM	Online Mendelian Inheritance in Man
pA	Polyadenylation
PCR	Polymerase chain reaction
PSD	Postsynaptic density
QMEAN	Qualitative Model Energy Analysis
que	*quetschkommode* mutant
RNA	Ribonucleic acid
RT-PCR	Reverse transcription polymerase chain reaction
SA	Splice acceptor
SCOP	Structural Classification of Proteins database
SDS	Sodium dodecyl sulphate
SERT	Serotonin transporter
sla	*schlaffi* mutant
SLC6A5	Solute carrier family 6 member 5 gene
SMO	Splice-site morpholino
SNVs	Single nucleotide variants
TAE	Tris-acetate EDTA
Tris-EDTA	Tris(hydroxymethyl)methylamine-ethylendiamine-tetraacetic acid
TM	Transmembrane domain
TMO	Translation-blocking morpholino
UAS	Upstream activation sequence
UCSC	University of California Santa Cruz
VIAAT	Vesicular inhibitory amino acid transporter